

## **2012 年度 修士論文**

# **表情・動作情報とタッチパネル操作情報の統合による 学習状態の推定**

提出日： 2013 年 2 月 1 日

指導： 山名 早人 教授

早稲田大学大学院 基幹理工学研究科 情報理工学専攻

学籍番号：5111B058-0

**須藤 優介**

## 概 要

効率的な学習指導を行うにあたり、生徒の学習状態を知ることは重要である。しかし、教師一人に対し多数の生徒が存在する状況や、家庭学習のような環境では、生徒一人ひとりの学習状態を把握することは難しい。このような問題を解決するために、学習状態を機械的に予測する研究が行われている。従来、学習状態推定の研究分野では、デスクトップ環境を前提とした研究が広く行われてきた。デスクトップ PC に対するキーボードやマウスによる入力、Web カメラにより取得される表情や動作、あるいは特殊なガジェットから得られる生体情報などは、学習状態を推定する上で有用な情報になる。一方で、教育現場ではタブレットデバイスの導入を積極的に進めており、今後タブレットデバイスを用いた学習状態の推定は重要度を増していくと考えられる。そこで本研究では、タブレットデバイスから得られる情報を用いた学習状態の推定手法を提案する。提案手法では、タブレットデバイスのカメラから得られる学習者の表情・動作と、ユーザの入力を特徴量として、「問題を解くのに悩んだか否か」、「問題に対し自信を持って答えたか否か」を推定した。また、デスクトップ環境と比べて学習者の表情や動作が安定して取得できない状況を考慮し、フェイストラッキングの処理状態に応じて動的に特徴量の重みを変化させた。これにより、解答時に悩んだか否か、自信があったか否かの推定について、それぞれ平均精度 75.3%, 66.7%, 平均再現率 74.8%, 60.1%を達成した。

# 目 次

第1章	はじめに.....	1
第2章	関連研究.....	2
2.1	概要 .....	2
2.2	表情・動作・生体情報に基づく学習状態推定 .....	2
2.3	端末操作情報に基づく学習状態推定 .....	2
2.4	表情・動作と端末操作情報に基づく学習状態推定 .....	4
2.5	まとめ.....	5
第3章	提案手法.....	6
3.1	概要 .....	6
3.2	対象とする学習環境.....	6
3.2.1	物理的学習環境 .....	6
3.2.2	学習教材.....	7
3.3	推定する学習状態 .....	8
3.4	表情・動作情報の解析 .....	8
3.5	端末操作情報の解析.....	10
3.6	機械学習を用いた学習状態の推定.....	10
第4章	実験・評価 .....	12
4.1	実験概要 .....	12
4.2	実験環境 .....	12
4.3	実験結果 .....	12
4.3.1	取得データの傾向.....	12
4.3.2	各手法との比較結果 .....	14
4.3.3	MKLによる各特徴量に対する重み付けの結果 .....	17
4.3.4	個人差に関する実験 .....	20
4.3.5	傾向の似たユーザで学習を行う手法の検討 .....	25
4.3.6	学習状態の表出が弱いデータを除去した実験.....	28
4.4	総合評価 .....	31
4.4.1	タブレットデバイスのみを用いた学習状態推定に関する評価 .....	32
4.4.2	表情・動作が安定して取得できないことを考慮した手法に関する評価 .....	33
4.4.3	個人差に関する評価 .....	33
4.5	実験・評価のまとめ.....	34
第5章	おわりに.....	37
5.1	本研究のまとめ.....	37
5.2	今後の課題.....	37

# 第1章 はじめに

効率的な学習指導を行うためには、教師が生徒の学習状態を理解し、学習状態に応じて指導方針を決定していくことが望ましい。しかし、一斉授業や家庭学習といった学習環境では、生徒一人ひとりの学習状態を教師が把握することは難しい。このような問題を解決するために、電子機器を用いて自動的に学習状態を推定するような研究が行われている[1]-[7]。

従来、e-Learning や一斉授業で使用される電子機器としては、デスクトップ型の PC が主流であった。そのため、学習状態推定の研究分野でも、デスクトップ型の PC で学習を行うユーザを対象とした研究が広く行われてきた。これらの研究では、生体情報、表情、動作、端末操作など、学習者の様々な情報を基に感情推定を行っている。学習者の感情を推定することにより、理解度合はどの程度か、教材が適切か否か、今後どのように教育していくかといった事の指標とすることができる。しかしながら、

- ・ デスクトップ PC を用いた授業自体が少ない
- ・ 特殊な装置を用いて生体情報を取得する、一般的でない学習教材を利用するなど、一般的でない学習環境を要する研究が多い

という問題があり、一般的な学習環境への適用は難しかった。

一方で、近年、教育現場ではタブレットデバイスの導入が進んでおり、総務省が展開する「フューチャースクール推進事業[8]」の中でも、すべての児童・生徒に対し1台ずつタブレットデバイスを配布することを目指している。教育現場のタブレットデバイス導入を受けて、タブレットデバイスを使用する学習者の学習状態推定は今後重要度を増していくと考えられるが、タブレットデバイスを用いて学習状態を推定するような研究は、我々の知る限り存在しない。タブレットデバイスを使用した学習状態推定を行い、その可能性・課題を評価することは、現在非常に重要であると考えられる。

そこで本研究では、タブレットデバイスから得られる情報のみを基に学習状態の推定を行う手法を提案する。本手法では、既存研究で学習状態推定への有用性が示されている「学習者の表情・動作」と「端末操作情報」をタブレットデバイスから取得し、それらを統合して学習状態を推定する。さらに、表情・動作情報を安定して取得することは困難であるという前提のもと、表情・動作情報取得状態に応じて、学習状態推定時の特徴量を選択する。学習環境としては、英語の四択問題に解答する環境を想定し、「悩んでいるか否か」、「自信を持って答えたか否か」の2つに関して推定を行う。

本稿は以下の構成をとる。まず2章では関連研究をまとめ、3章で提案手法について説明する。さらに4章で実験・評価について述べ、5章でまとめを行う。

## 第2章 関連研究

### 2.1 概要

タブレットデバイスを用いた学習状態推定の研究は、我々の知る限り存在しない。本章では、デスクトップ PC で学習する学習者を対象とした、既存の学習状態推定に関する研究について述べる。

### 2.2 表情・動作・生体情報に基づく学習状態推定

学習者の表情、動作、生体情報を基に感情推定を行った研究について、表 1 にまとめた。

これらの研究では、表情・動作、生体情報といった情報が、学習状態推定に有用な情報であることが示されている。

しかし、表情や動作に基づく学習状態推定の研究では、表情や動作が取得できない状態では推定が行えないという問題がある。繁田ら[2]の研究では、初期フレームにおいて顔の特徴点を手動で設定するにも関わらず、フェイストラッキング中に特徴点が 5 画素以上ずれる割合が 58~89%になるという結果が出ている。学習中の表情の変化は、笑顔などと比べると小さな物であり、特徴点が正確に取得できていない状態では十分な精度を出すことは難しい。繁田らの実験の中では高い再現率での推定に成功しているが、テスト用の被験者が 3 名と少なく、実験の中で、顔特徴が完全に取得できないようなケースは想定されていない。

また、生体情報を取得する研究[1][4]では、生体情報取得のために学習者が特殊な装置を装着する必要がある上に、学習者の負担が大きくなる上に、一般的な学習環境下では適用するのが難しい。中村ら[3]の研究では、顔特徴の取得に 2 台のカメラを使用しており、こちらも一般的な学習環境への導入は難しい。

### 2.3 端末操作情報に基づく学習状態推定

端末操作情報を用いて学習状態推定を行った中村ら[5]の研究について、表 2 にまとめた。中村らの研究では、操作時間間隔を基に、学習者の行き詰まり状態を約 80%の再現率で検出している。

しかし、学習者自身の操作で演習問題に対するヒントを自由に表示・閲覧できるような特殊な環境下で実験を行っており、一般的な学習環境下での有用性は示されていない。

表 1 表情・動作・生体情報に基づく学習状態推定の研究

年	著者	使用特徴量	推定項目	学習環境	推定手法	性能
‘07	黒川ら [1]	・ 眉毛の動き ・ 唇の動き ・ 脈拍数 ・ 呼吸数 ・ 皮膚温度	① 理解している/ していない ② 飽きている/ 興味がある ③ 疲れている/ 集中している ④ 難しい/簡単	・ 実際に使用されている情報通信に関する教材 ・ 生体信号測定装置と家庭用カメラを使用	テンプレートマッチング法とマハラノビスの距離を用いたアルゴリズム	再現率 ① 69% / 75% ② 73% / 0% ③ 71% / 56% ④ 57% / 82% ※学習に10人, テストに5人の男子大学生
‘09	繁田ら [2]	・ 眉毛の動き ・ 唇の動き	① 理解している/ していない ② 飽きている/ 興味がある ③ 疲れている/ 集中している ④ 難しい/簡単	・ 実際に使用されている通信工学入門と, 経済学入門の教材 ・ ビデオカメラで被験者を撮影	・ 顔特徴点平均変化率の大小を基に被験者を2つのグループに大別 ・ 推定に用いる特徴点を感情対ごとに選択 ・ [1]と同様のアルゴリズム	再現率 ① 81.3% / 33.3% ② 100% / 83.3% ③ 100% / 55.6% ④ 83.3% / 80.0% ※学習に23人(男18, 女5), テストに3人(男2, 女1)の大学生
‘10	中村ら [3]	・ 顔の動き ・ 顔の傾き ・ 凝視時間 ・ つぶやき時間	・ 主観的難易度	・ 四則演算の穴埋め問題 ・ ディスプレイ上に2台のカメラを設置して被験者を撮影	・ 学習者ごとに各特徴量の重みを学習 ・ SVMを用いて識別	精度 ・ 入力要求型学習環境で平均 86.2% ・ 閲覧主体型学習環境で74.9% ※21歳から33歳の男女11人でクロスバリデーション
‘12	Szafirら [4]	・ 脳波	・ 集中力の低下	・ ロボットから話を聞き, 後でどれだけ記憶しているかを問う質問に答える ・ 脳波測定ヘッドセットを装着	・ 取得したデータから閾値を算出する	記憶力が43%増加 ※被験者30人(男女各15人) ※推定精度に関する議論は無い

表 2 端末操作情報に基づく学習状態推定の研究

年	著者	使用特徴量	推定項目	学習環境	推定手法	性能
'02	中村ら [5]	操作時間間隔	行き詰まり状態	<ul style="list-style-type: none"> <li>・ Java 言語 穴埋め問題</li> <li>・ 学習者自身の操作で演習問題に対するヒントを自由に表示・閲覧できるような支援機能のある環境</li> </ul>	実験的に閾値を設ける	再現率 約80% ※男子大学生20名

## 2.4 表情・動作と端末操作情報に基づく学習状態推定

表情・動作情報と端末操作情報を両方用いた学習状態推定の研究について、表 3 にまとめた。Kapoor ら[6]の研究では、興味があるか否かについて約 86%の精度、中村ら[7]の研究では、主観的難易度について約 85%の精度で推定に成功している。しかし、Kapoor らの手法では赤外線カメラと圧力センサ付きの椅子、中村らの研究ではステレオカメラを使用するなど、両者ともに特殊な装置を用いて学習者の情報を取得している。これらの装置はいずれも一般的な学習環境では導入が難しい。

表 3 表情・動作と端末情報に基づく学習状態推定の研究

年	著者	使用特徴量	推定項目	学習環境	推定手法	性能
'05	Kapoor ら [6]	① 眉毛・目・口の動き ② 頭の動き ③ 姿勢 ④ 問題のレベル・進行状況	興味があるか否か	<ul style="list-style-type: none"> <li>・ 教育用パズルゲーム（難易度が事前に分かっている）</li> <li>・ LEDカメラや圧力センサがついた椅子を使用</li> </ul>	混合ガウス過程を用いて、各特徴量の重みを計算して推定を行う	Accuracy 86%以上 ※8歳から11歳の子供8人、クロスバリデーション
'07	中村ら[7]	<ul style="list-style-type: none"> <li>・ 表情</li> <li>・ 操作時間間隔</li> </ul>	主観的難易度	<ul style="list-style-type: none"> <li>・ 4択式英単語問題</li> <li>・ ディスプレイ上に2台のカメラを設置して被験者を撮影</li> </ul>	<ul style="list-style-type: none"> <li>・ 学習者ごとに各特徴量の重みを学習</li> <li>・ SVMを用いて識別</li> </ul>	精度 平均85%前後 ※被験者4名

## 2.5 まとめ

本章では，既存の学習状態推定に関する研究として，

1. 表情・動作や生体情報を使用した研究
2. 端末操作情報を使用した研究
3. 表情・動作と端末操作情報を組み合わせた研究

を紹介した．

いずれも高い精度での推定に成功しているが，

- ・ 特殊な装置を用いている
- ・ 特定の学習用ソフトウェアでしか実験を行っていない
- ・ 表情・動作を取得できない環境では推定を行えない

といった問題が残されている．既存手法の問題点について，表 4 にまとめた．

表 4 既存研究の問題点

年	著者	問題点
‘07	黒川ら[1]	・ 生体情報を取得するのに特殊な装置が必要
‘09	繁田ら[2]	・ 表情が取得できない状況では推定が行えない
‘10	中村ら[3]	・ 表情が取得できない状況では推定が行えない ・ 2台のカメラが必要
‘12	Szafir ら[4]	・ 生体情報を取得するのに特殊な装置が必要
‘02	中村ら[5]	・ 特殊な学習教材が必要
‘05	Kapoor ら[6]	・ 赤外線カメラ，圧力センサ付きの椅子など，特殊な装置が必要 ・ 問題の難易度が事前に分かっている必要がある
‘07	中村ら[7]	・ 2台のカメラが必要



## 第3章 提案手法

### 3.1 概要

提案手法では、タブレットデバイスを用いた学習状態推定という今までに無い試みを行うと同時に、既存研究の問題点解決を図る。具体的には、以下の様な要件を満たす学習状態推定システムを構築する。

- ・ 特殊な装置を用いない
- ・ 学習教材に対する制限が弱い
- ・ 表情・動作が安定して取得できないことを考慮した学習状態推定

本章では、以上の要件を満たすことを念頭に置き、本研究で対象とする学習環境、推定する学習状態、提案手法の処理手順について述べる。

### 3.2 対象とする学習環境

提案手法では、学習環境において以下の2つの要件を満たす必要がある。

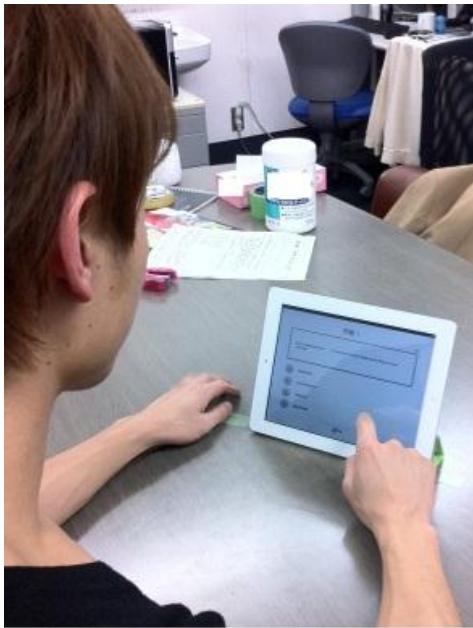
- ・ 特殊な装置を用いない
- ・ 学習教材に対する制限が弱い

ハードウェア（使用する装置）、ソフトウェア（学習用教材）の2つの観点から、対象とする学習環境について述べる。

#### 3.2.1 物理的学習環境

提案手法では、タブレットデバイス1台のみを使用して学習状態の推定を行う。実際に実験を行った時の様子を、図1(a)に示す。タブレットデバイスは、学習者の顔が内側カメラに写る位置に設置し、角度をつけた状態で固定する。

学習者には「タブレットデバイスを動かさない」という制限を設けることになるが、タブレットデバイス1台で構築可能な環境であり、特殊な装置を用いないという要件は満たしている。



(a)



(b)

図 1 (a)解答中の様子 (b)実際の解答画面

### 3.2.2 学習教材

端末操作情報を使用する研究[5][6][7]では、学習教材に制限を加えることで有用な特徴量を抽出しているものがある。中村ら[5]の研究では、学習者自身の操作で演習問題に対するヒントを自由に表示・閲覧できるような環境の下、端末操作時間間隔を利用して学習状態の推定を行なっている。Kapoor ら[6]の研究では、事前に難易度が分かっているクイズ問題を使用し、解いている問題の難易度と進行状況の特徴量としている。

提案手法では、タブレット端末を用いて英語の四択問題に解答するような環境を対象とする。これは、中村ら[7]らが用いた「英単語の4択問題」という学習教材に近い。実際の解答画面を図 1(b)に示す。

選択式の問題は、センター試験でも採用されるなど広く使用されており、学習教材・問題としては一般的なものであると考えられる。また、選択式の問題では解答の途中経過が見られず、学習者の理解度を解答結果から推測することが難しい。これらの理由から、選択式問題は、学習状態の推定が重要な意味を持つ学習環境であると考えた。

多少の制限はあるものの、一般的に広く使用されている問題形式であり、

- ・ 学習環境（教材，問題）に対する制限が弱い

という要件は満たしていると考えられる。

### 3.3 推定する学習状態

提案手法では、推定する学習状態を、

- ・ 悩んだか否か
- ・ 自信を持って答えたか否か

の2つとする。

本研究では、「教師が生徒の理解度を正確に把握すること」を目指している。選択式問題への解答では、解答までの過程が不可視であるため、

- ・ 答えに行き着くまでにどれだけ深く考えたのか
- ・ 最終的に根拠を持って解答したのか

が分からない。このような背景から、「悩んだか否か」、「自信を持って答えたか否か」という学習状態は、教師が生徒の理解度を把握する上で有用な判断材料になると考えた。

### 3.4 表情・動作情報の解析

表情・動作情報の取得には、Active Shape Model[9]を用いる。モデルベースの手法は、最新の感情推定研究[10]でも有用性が示されており、カメラ以外の特殊な機器を必要としないため、提案手法での採用を決定した。なお、提案手法では、Yao Wei らが開発したASMLibrary[11]を使用する。

Active Shape Model により、動画像上の顔領域から図 2 のように 68 点の特徴点を得ることができる。特徴点は、動画像の左上の点を原点として、画像右方向を  $x$  軸正方向、下方向を  $y$  軸正方向とした  $xy$  平面上の座標で表現される。このように取得した特徴点を基に、以下の特徴量を抽出する。

- A) 眉毛の動き
- B) 口の動き
- C) 目の開閉の動き
- D) 頭の動き（鼻の座標の動き）
- E) 画面からの距離（検出した顔の大きさ）
- F) 顔の傾き
- G) 特徴点検出に成功したフレームの割合

※68 点すべての特徴点に成功した状態を成功とする

特徴量の選定は既存研究を基に行った。表情の変化は、眉毛・目・口の動きによってなされ、既存手法でも表情の変化を取得するためにこれらの特徴量として使用している。提案手法では、眉毛・目・口の動きを、特定の特徴点同士の距離の変化で表現する。例として、提案手法における眉毛の動きの取得方法を以下に示す。

step1. 動画の  $N$  フレーム目で、目の上部の点と、眉毛上の 3 点を繋いだ 3 本の線分

の長さを合計する（左右それぞれについて行う）

step2. step1 で算出した値を，検出した顔の大きさで正規化する（左右それぞれについて行う）

step3. N+1 フレームでも，step1, step2 と同じ操作を行う

step4. step2 の算出結果と，step3 の算出結果の差分を取る（左右それぞれについて行う）

step5. 左右それぞれの差分を合計する

step6. 全フレームに対して step1~step5 の処理を行い，step5 の算出結果の総和をとる

step7. 処理したフレーム数で step6 の結果を割る（正規化）

口の動き，目の動きに関しても，上記の処理手順と同様にして取得する．

動作情報としては，頭の動き，画面からの距離，頭の傾き，特徴点の検出に成功したフレーム数を取得する．前述の眉毛や目の動きと同じように，頭の動きや画面からの距離，顔の傾きも，前のフレームとの差分の総和により表現する．また，学習者の動きが大きいとトラッキングに失敗することが多いため，特徴点検出に成功したフレームの割合を，動きの大きさを表す特徴量として使用している．

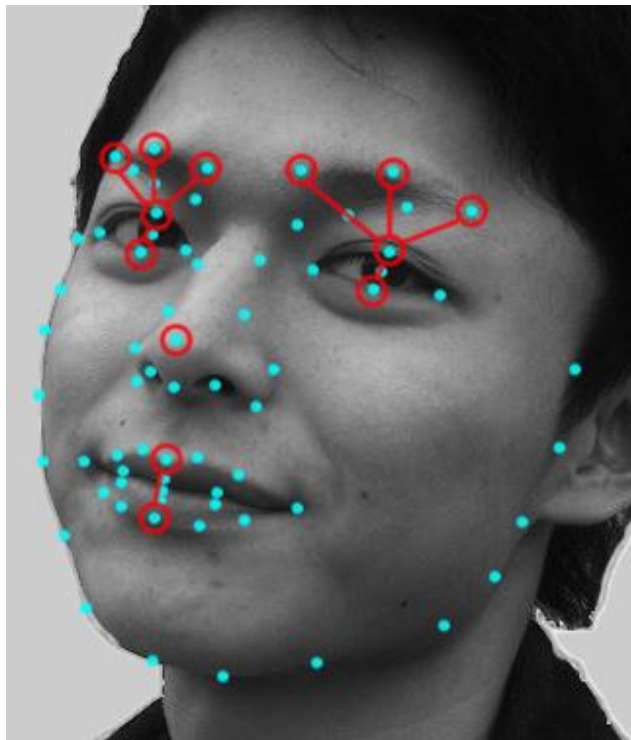


図 2 Active Shape Model によるアライメント

※ 画像は著者自身

### 3.5 端末操作情報の解析

提案手法では、端末操作情報として、以下の特徴量を抽出する。

- a) 答えがあっているか否か
- b) 問題に答えるのにかった時間
- c) 答えを変更した回数
- d) 最後に答えを選んでから次の問題へ行くまでの時間
- e) 最後の答えを選択した時のタップ時間（画面に触れていた時間）
- f) タップしていた時間（端末に触れていた時間）の合計

なお、本研究では、被験者に以下の手順で問題に解答してもらい、上記の特徴量を抽出する。

*step1.* 解答時間、問題の形式、解答方法の説明を受ける

*step2.* 準備が出来たら、「START」ボタンを押し、解答を開始する

*step3.* 解答を選択し、その解答で決定ならば「次へ」ボタンを押す

※ 「次へ」ボタンを押すと前の問題に戻ることは出来ない

*step4.* 30 問の問題に答え終わるまで解答を続ける

端末操作情報としては、操作時間間隔が学習状態推定に有用であることが既存研究で示されている[5][7]。本研究では、英語の四択問題に解答する環境を対象として学習状態の推定を行うため、操作時間間隔を細分化し、「問題に答えるのにかった時間」と「最後に答えを選んでから次の問題へ行くまでの時間」を取得している。また、答えの変更回数も、「悩んだか否か」の推定に有効な特徴量であると判断し、抽出を行なっている。

さらに、本研究ではタッチデバイスで操作情報を取得するため、タッチデバイス特有の情報として、画面に触れている時間の長さを取得している。

### 3.6 機械学習を用いた学習状態の推定

提案手法では、Multiple Kernel Learning(MKL)[12]を用いて、表情・動作情報と端末操作情報を統合した学習状態推定を行う。MKLを使用することにより、各特徴料に対する最適な重み付けを学習して、学習状態を推定することが可能になる。

また、提案手法では、表情・動作情報が安定して取得できないことを前提として学習状態推定を行う。具体的には、動画像処理において特徴点検出に成功しているフレームの割合が一定以下だった場合、端末操作情報のみを用いた推定に切り替えることで、推定精度・再現率の向上を図る。なお、提案手法では、経験的に、特徴点検出に成功しているフレームが全フレームの3割未満だった場合に、端末操作情報のみを用いた推定に切り替える。

提案手法における処理の流れについて図 3 にまとめた。

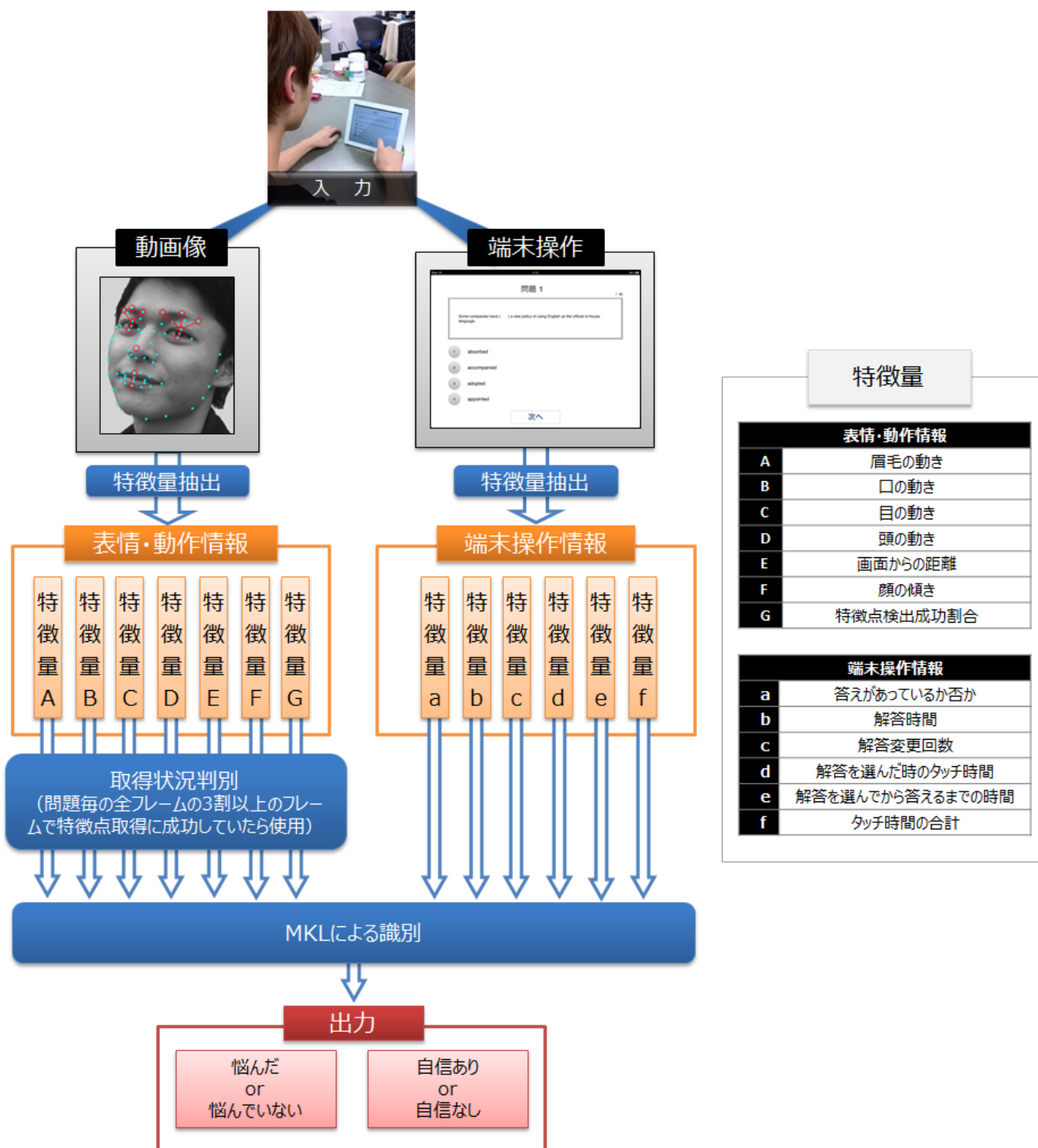


図 3 提案手法における処理の流れ

## 第4章 実験・評価

### 4.1 実験概要

提案手法の評価を行うため、評価実験を行った。本章では、実験の方法と結果、および評価について述べる。

### 4.2 実験環境

本実験におけるデータ取得環境について、表 5 にまとめた。被験者には、タブレットデバイス上で全ての問題に解答してもらった後、各問題で悩んだか否か、自信を持って答えたか否かを問うアンケート（付録 B）に答えてもらった。

表 5 データ取得環境

被験者	大学生 25 名（男 11：女 14）
問題	英語の 4 択問題 10 問 （センター試験レベル 20 問と中学英語レベル 10 問（付録 A）） 参考：図 1(b)
制限時間	30 分
デバイス 設置環境	角度をつけて固定 参考：図 1(a)
その他	<ul style="list-style-type: none"><li>一度解答すると前の問題には戻れない キャプチャを行うことは事前に伝えている</li><li>「カメラに写る位置から動かないように」など、学習者の動きに対して制限は設けていない</li><li>全員に対し同じ問題を出題</li></ul>

### 4.3 実験結果

#### 4.3.1 取得データの傾向

我々が取得したデータの全体的な傾向について述べる。

表 6 に、アンケート集計結果についてまとめた。25 人の被験者に 30 問の問題を解いてもらったため、全 750 の回答を得ている。なお、本実験では、「まったく自信がなかった」、「あ

まり自信がなかった」という回答を「自信がなかった」としてまとめ、「少し自信があった」、「自信があった」という回答を「自信があった」としてまとめている。同様に、「とても悩んだ」、「少し悩んだ」を「悩んだ」、「あまり悩まなかった」、「まったく悩まなかった」を「悩まなかった」としてまとめている。

全体としては、悩んだ問題、悩まなかった問題の数はそれぞれほぼ半数となった。一方で、自信があったという回答が、自信がなかったという回答の数を大きく上回っており、データ取得数に若干の偏りが生じる結果となった。また、集計結果から、

- ・ 答えに自信が無い場合は多くの場合悩んでいる
- ・ 悩んだ末、自信を持って答えるケースが多く存在する

ということもわかった。また、問題毎・学習者毎の結果は付録 C にまとめた。

表 6 アンケート集計結果

	自信がなかった	自信があった	合計
悩んだ	252	100	352
悩まなかった	28	370	398
合計	280	470	750

解答時間の傾向について表 7 にまとめる。すべて 4 択問題ではあるが、解答時間に大きくばらつきがあることが見てとれる。

表 7 解答時間の傾向

平均解答時間	19.2 sec
標準偏差	13.5
最長解答時間	130 sec
最短解答時間	2.8 sec

さらに、その他の傾向について表 8 にまとめる。ここで、特徴点取得の成功・不成功の判断は、各被験者に対し問題毎に行なっている。1 フレームも顔特徴点を取得できていなかったのべ問題数が 19%に達しており、顔特徴点が取得できない状況を考えることには大きな意味があると言える。また、問題毎・学習者毎の正答率は付録 C にまとめた。

表 8 取得したデータの傾向

正答率	69.2 %
顔特徴点平均取得率 $\left( \frac{\text{取得成功フレーム}}{\text{全フレーム}} \times 100 \right)$	67.1 %
当該問題中、3 割以上のフレームで顔特徴点取得に成功していたのべ問題数の割合	72 %
当該問題中、1 フレームも顔特徴点取得に成功していなかったのべ問題数の割合	19.2 %



### 4.3.2 各手法との比較結果

#### a) 各手法との比較実験の概要

提案手法が有効に機能しているか調べるために、以下の3パターンとの比較を行った。

- ① 全データ（表情・動作，端末操作情報のすべて）を使用し，MKLを用いて分類する（表情・動作情報の取得状態に応じた使用特徴量の切り替えを行わない）
- ② 端末操作情報のみを利用し，MKLを用いて分類する
- ③ 全データ（表情・動作，端末操作情報のすべて）を使用するが，特徴量の重み付けをせずに分類する

比較実験では，25人分のデータを5人ずつの5セットに分割し，leave-one-out cross-validationにより，Precision, Recall, F値を算出している．なお，本実験では，正解セットとしてアンケート結果を使用する．

#### b) 各手法との比較実験の結果

以下，図4，図5，図6に，各手法を用いて各学習状態の推定を行った時のPrecision, Recall, F値を示す．

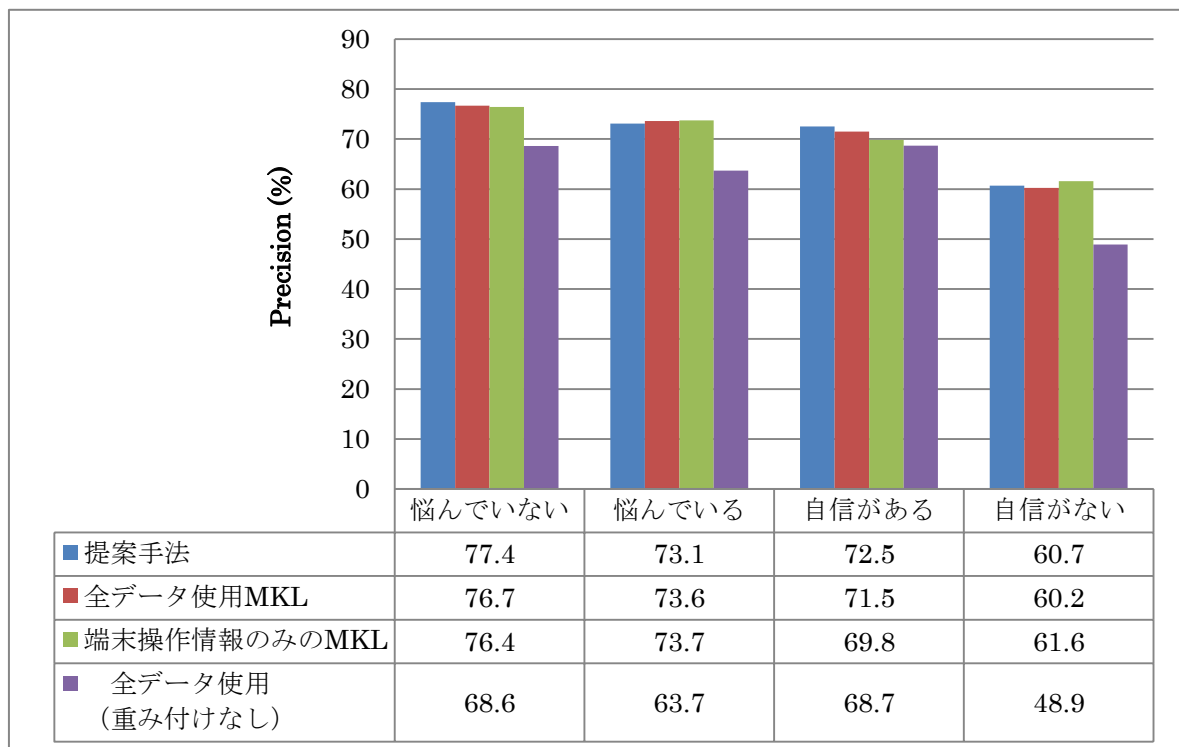


図4 実験結果 (Precision)

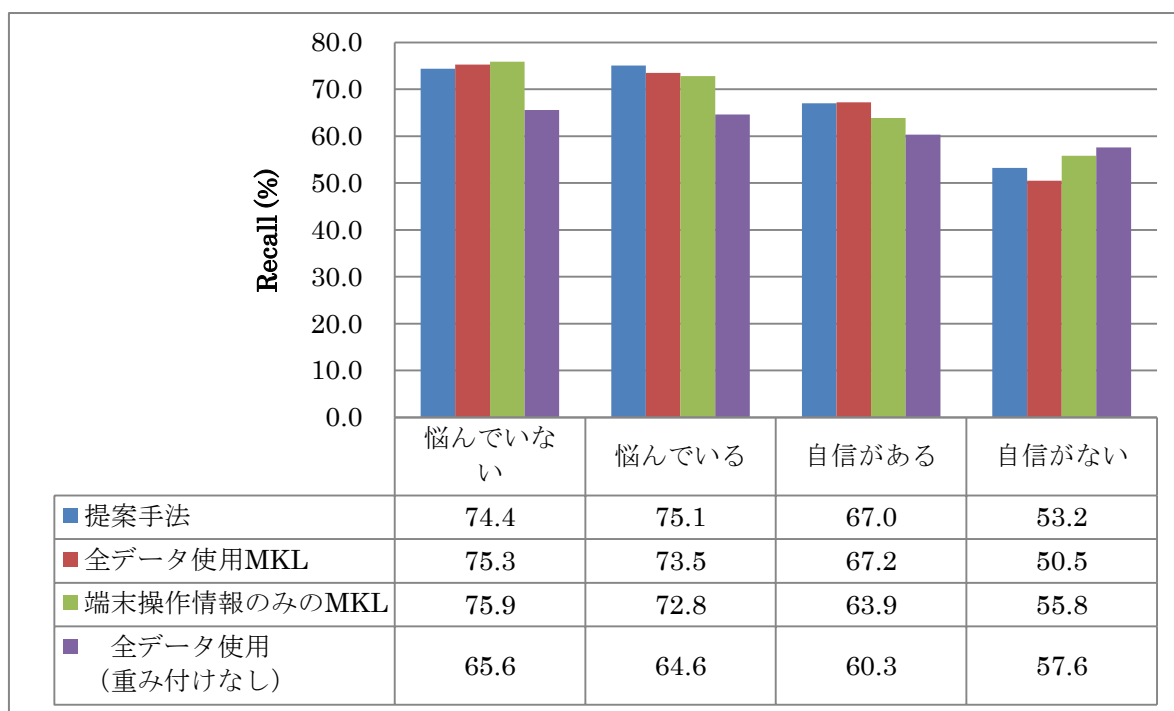


図 5 実験結果 (Recall)

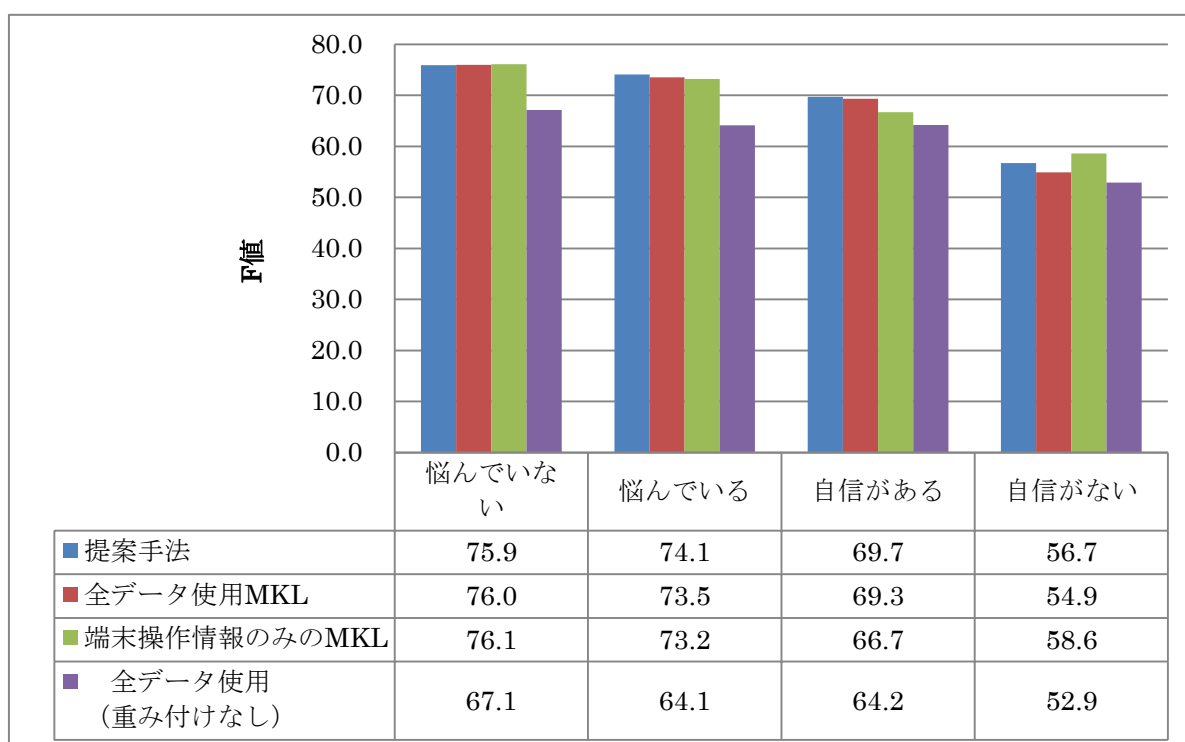


図 6 実験結果 (F 値)

### c) 各手法との比較実験の評価

実験結果 図 4, 図 5, 図 6 を見ると, Recall 以外では「全データ使用 (重み付けなし)」の結果を提案手法が大きく上回っているのが分かる. これは MKL による各特徴量への重み付けが有効に働いた結果であるといえる.

一方で, 「全データを使用し, MKL を用いて分類」した結果と, 「端末操作情報のみを利用し, MKL を用いて分類」した結果には, 提案手法との間に大きな差が見られない. さらなる尺度での比較を行うため, 3 つの手法について, Accuracy の算出も行った. 結果を図 7 に示す.

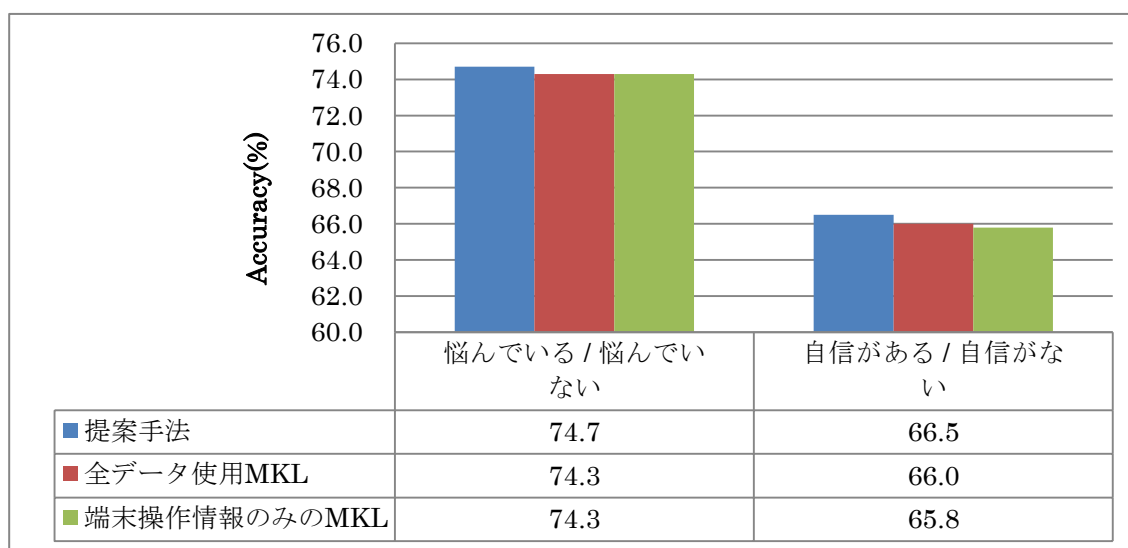


図 7 実験結果 (Accuracy)

Accuracy の比較では, 他の手法に比べ提案手法の方が僅かに優れた数値を記録している. しかしながら, 他の実験結果と同様, その差は僅差である.

各実験結果において, 提案手法と他の手法の間に統計的に有意な差が存在するのかを調べるために, 検定を行った. 検定に用いたデータは, 25 名の各被験者に対する各手法の適用結果である. 母集団の分布形態が不明であるため, 検定手法としてマン・ホイットニーの U 検定を用いた. Precision, Recall, F 値, Accuracy に対する各検定結果 (P 値) を, 表 9 ~ 表 12 に示す. なお, 網掛けは, その値が 0.05 以下であることを表す.

表 9 提案手法と各手法の Precision に関する検定結果 (P 値)

	悩んでいる	悩んでいない	自信がある	自信がない
全データ使用 (重み付けあり)	0.45	0.40	0.47	0.46
端末操作情報使用	0.50	0.37	0.34	0.46
全データ使用 (重み付けなし)	0.019	0.039	0.33	0.14

表 10 提案手法と各手法の Recall に関する検定結果 (P 値)

	悩んでいる	悩んでいない	自信がある	自信がない
全データ使用(重み付けあり)	0.37	0.43	0.40	0.25
端末操作情報使用	0.33	0.38	0.21	0.37
全データ使用(重み付けなし)	0.15	0.18	0.36	0.33

表 11 提案手法と各手法の F 値に関する検定結果 (P 値)

	悩んでいる	悩んでいない	自信がある	自信がない
全データ使用(重み付けあり)	0.43	0.48	0.44	0.25
端末操作情報使用	0.38	0.43	0.22	0.38
全データ使用(重み付けなし)	0.012	0.024	0.38	0.34

表 12 提案手法と各手法の Accuracy の差に関する検定結果 (P 値)

	悩んでいるか否か	自信があるか否か
全データ使用	0.44	0.43
端末操作情報使用	0.42	0.38

表 9～表 12 を見ると、有意水準 5%において提案手法の優位性が示されたのは、重み付けをせずに全データを使用した場合に対する Precision, F 値のみであった。すなわち、

- ① 全データ（表情・動作、端末操作情報のすべて）を使用し、MKL を用いて分類する（表情・動作情報の取得状態に応じた使用特徴量の切り替えを行わない）
  - ② 端末操作情報のみを利用し、MKL を用いて分類する
- の 2 つの手法に対する提案手法の優位性は確認されなかった。

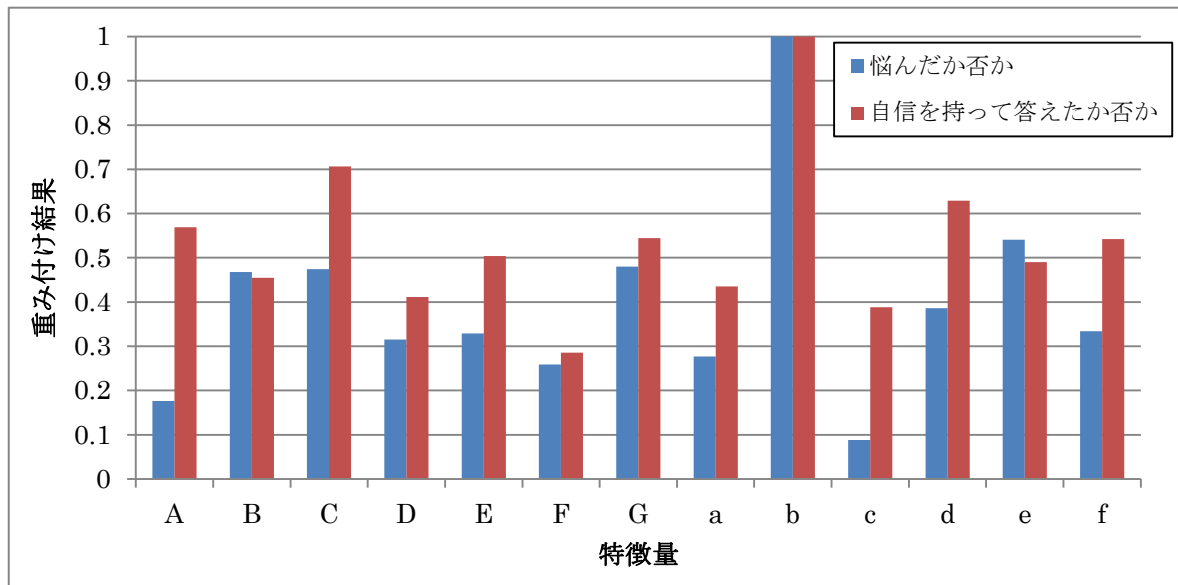
### 4.3.3 MKL による各特徴量に対する重み付けの結果

MKL による各特徴量に対する重み付けの結果を、図 8 に示す。

#### a) 悩んだか否か推定時の各特徴量の重み付けに関する考察

“悩んだか否か” 推定時の重みが高い順に特徴量を並べると、以下のようになる。（重みが 0.3 以下のものは表記しない）

- |                    |                 |
|--------------------|-----------------|
| ① 解答時間             | ⑤ 口の動き          |
| ② 解答を選んでから答えるまでの時間 | ⑥ 解答を選んだ時のタッチ時間 |
| ③ 特徴点検出成功割合        | ⑦ タッチ時間の合計      |
| ④ 目の動き             | ⑧ 画面からの距離       |
|                    | ⑨ 頭の傾き          |



表情・動作情報	
A	眉毛の動き
B	口の動き
C	目の動き
D	頭の動き
E	画面からの距離
F	顔の傾き
G	特徴点検出成功割合

端末操作情報	
a	答えがあっているか否か
b	解答時間
c	解答変更回数
d	解答を選んだ時のタッチ時間
e	解答を選んでから答えるまでの時間
f	タッチ時間の合計

図 8 各特徴量に対する重み付けの結果

#### a-1) 解答時間と解答を選んでから答えるまでの時間

「解答時間」、「解答を選んでから答えるまでの時間」は、「次へボタンを押すまでの時間」、「解答ボタンを押してから次へボタンを押すまでの時間」と言い換えられ、いずれも「操作時間間隔」に近い意味合いを持つ。操作時間間隔は、既存研究[5][7]の中で、“行き詰まり状態”や“主観的難易度”の推定に対する有効性が明らかにされている。今回の実験では学習用教材として4択問題を用いており、途中経過を書くようなことは無い。また、「次へ」ボタンを押したら答え直すことはできないという仕様上、被験者が「次へ」ボタンを押すのに慎重になることが考えられる。このような学習環境の性質上、今回、「解答時間」、「解答を選んでから答えるまでの時間」の重み付けが高くなったと考えられる。

### a-2) 体の動きと顔の動き

「特徴点検出成功割合」は、顔の動きが少ないほど特徴点が取れやすいという性質を考え、顔の動きの大きさを表す尺度として抽出している。すなわち、「特徴点検出成功割合」、「画面からの距離」、「頭の傾き」は、いずれも被験者の体の動きを表す特徴量であると言える。「体の動き」、「目の動き」、「口の動き」は、既存研究[1][2][3][6][7]の中でも学習状態推定への有効性が明らかになっている。今回の実験でも、その有効性が示される結果となった。

### a-3) タッチデバイス特有の特徴

その他に高い重み付けがされていたのが、「解答を選んだ時のタッチ時間」と「タッチ時間の合計」である。これらはタッチデバイス特有の特徴量であり、既存研究では使用されていない。提案手法による“悩んだか否か”の推定では、Accuracy：74.6%という高い値を記録している。学習環境に既存手法のような強い制限を加え無かったにもかかわらず高い値を出せたのは、タッチデバイス特有の特徴量がうまく機能し、MKLによる各特徴量に対する重み付けも効果的であったためと考えられる。

## b) 自信を持って答えたか否か推定時の各特徴量の重み付けに関する考察

“自信を持って答えたか否か”推定時の重みが高い順に特徴量を並べると、以下のようになる。（重みが0.3以下のものは表記しない）

- |                 |                    |
|-----------------|--------------------|
| ① 解答時間          | ⑦ 画面からの距離          |
| ② 目の動き          | ⑧ 解答を選んでから答えるまでの時間 |
| ③ 解答を選んだ時のタッチ時間 | ⑨ 口の動き             |
| ④ 眉毛の動き         | ⑩ 答えがあっているか否か      |
| ⑤ タッチ時間の合計      | ⑪ 頭の動き             |
| ⑥ 特徴点検出成功割合     | ⑫ 解答変更回数           |

### b-1) 全体的な重み付けの傾向に関する考察

「解答時間」が最も高く重み付けされている点は“悩んだか否か”の推定と同様だが、全体として、“悩んだか否か”の推定よりも特徴量の重み付けに偏りが無い。提案手法による“自信を持って答えたか否か”の推定は、Accuracy：66.5%であり、“悩んだか否か”の推定における74.6という数値を大きく下回っている。この推定結果と重み付けの偏りが少ない結果を見ると、今回使用した特徴量の中に、“自信を持って答えたか否か”を推定する上で際立って有用な特徴量が存在しなかったことが考えられる。あるいは、被験者毎に学習状態が表出する特徴量が異なっており、全体としての重み付けがばらけてしまった可能性も考えられる。これらを明らかにするには、推定結果の個人差に関する調査が必要である。

### b-2) タッチデバイス特有の特徴

提案手法における“自信を持って答えたか否か”の推定では、十分な精度・再現率を出せていなかったが、“悩んだか否か”の推定と同様に、ここでも「解答を選んだ時のタッチ時間」、

「タッチ時間の合計」が高く重み付けされている点に注目したい。「解答を選んだ時のタッチ時間」、「タッチ時間の合計」はともにタッチデバイス特有の情報であり、本研究で初めて採用した特徴量である。「悩んだか否か」、「自信を持って答えたか否か」の両方の推定でこれらの特徴量が高く重みづけされたことは、今後のタッチデバイスによる学習状態推定の研究にとって重要な結果である。

#### 4.3.4 個人差に関する実験

##### a) 個人差に関する実験の概要

2010 年中村ら[3]の研究では、学習中に現れる表情の変化には個人差があるという前提の下、学習者毎に識別器を作ることにより高い精度を実現している。本研究においても表情を使用して学習状態の推定を行っており、各特徴量の重要度や提案手法の効果に個人差があるか否か調べることは重要である。

本項では、「表情・動作情報を用いることで生じる推定結果の変化」、および、「表情・動作情報の取得状態に応じた特徴量の選択による推定結果の変化」に個人差があるのかを調べた結果について述べる。具体的には

- ① 提案手法
- ② 全データを使用し、MKL を用いて分類する手法
- ③ 端末操作情報のみを利用し、MKL を用いて分類する手法

の 3 パターンによる推定結果を個人毎に算出し、

- ・ ①の精度 - ②の精度
- ・ ①の精度 - ③の精度
- ・ ①の再現率 - ②の再現率
- ・ ①の再現率 - ③の再現率

を計算して、精度・再現率について箱ひげ図を作成した。

##### b) 個人差に関する実験の結果

以下、各実験結果について図にまとめる。

- ・ 図 9 : 提案手法と他の手法の、個人毎の推定精度の差分の箱ひげ図
- ・ 図 10 : 提案手法と他の手法の、個人毎の再現率の差分の箱ひげ図

なお、図 9、図 10 では、正の値が提案手法の方が優れていることを示し、負の値は提案手法のほうが劣っていることを示す。また、箱の大きいほど、ひげが長いほど、提案手法と他の手法の推定結果の差分に個人差があることを示す。

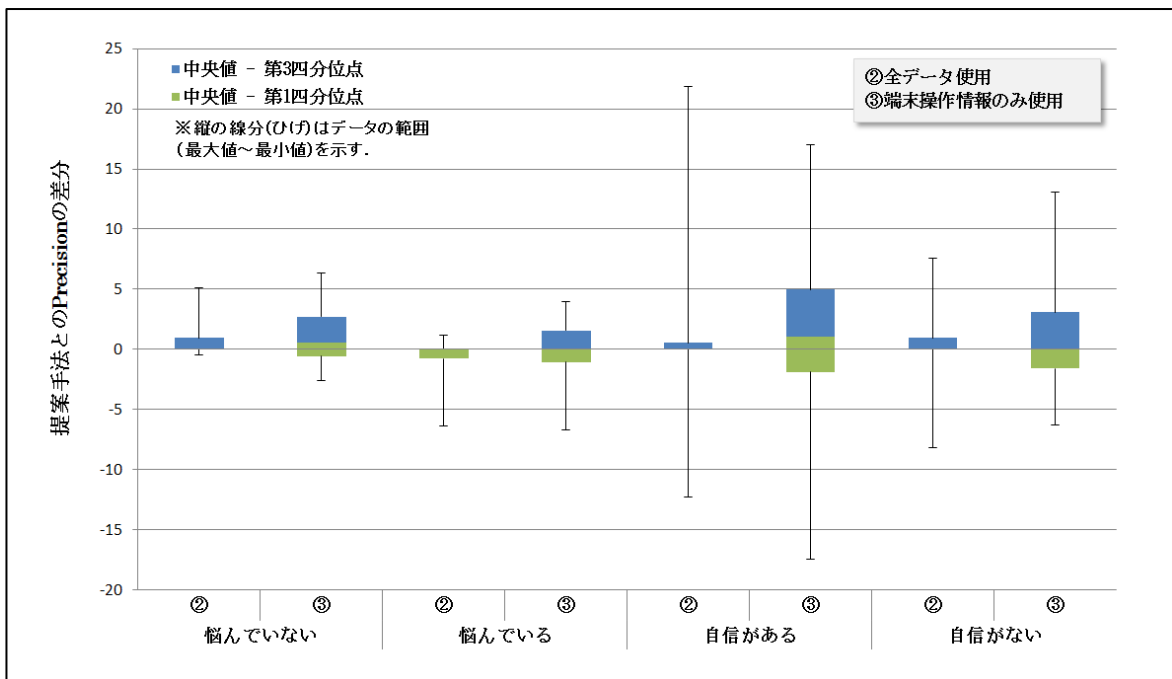


図 9 提案手法と他の手法の Precision の差分を示す箱ひげ図

正の値は提案手法の方が優れていることを示し、負の値は提案手法の方が劣っていることを示す。

箱とひげの大きさはデータのばらつきの大きさを示す。

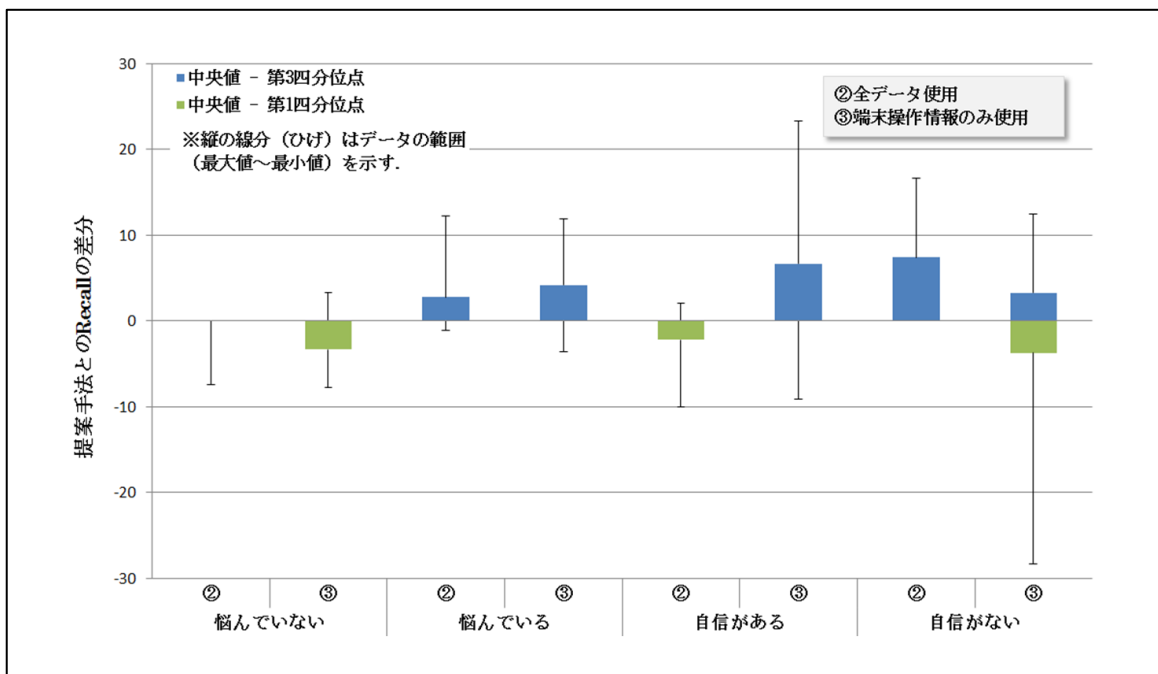


図 10 提案手法と他の手法の Recall の差分を示す箱ひげ図

正の値は提案手法の方が優れていることを示し、負の値は提案手法の方が劣っていることを示す。

箱とひげの大きさはデータのばらつきの大きさを示す。



実験結果，図 9，図 10 を見ると，学習者毎に提案手法の有効性に大きな差があることが見て取れる。

各学習者に各手法を適用したときどのような結果になっているのかを示すため，学習者ごとの Accuracy を算出した．悩んでいるか否かについて図 11，自信があるか否かについて図 12 に示す．

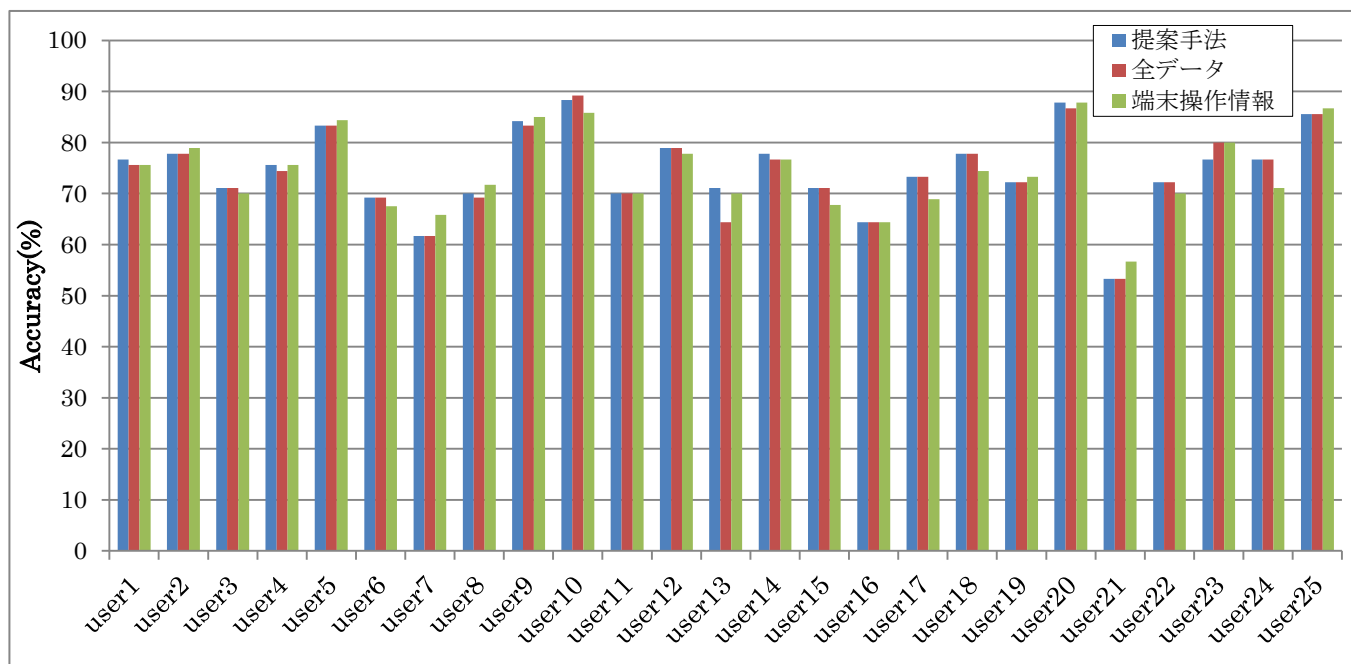


図 11 学習者毎の Accuracy (悩んでいるか否か)

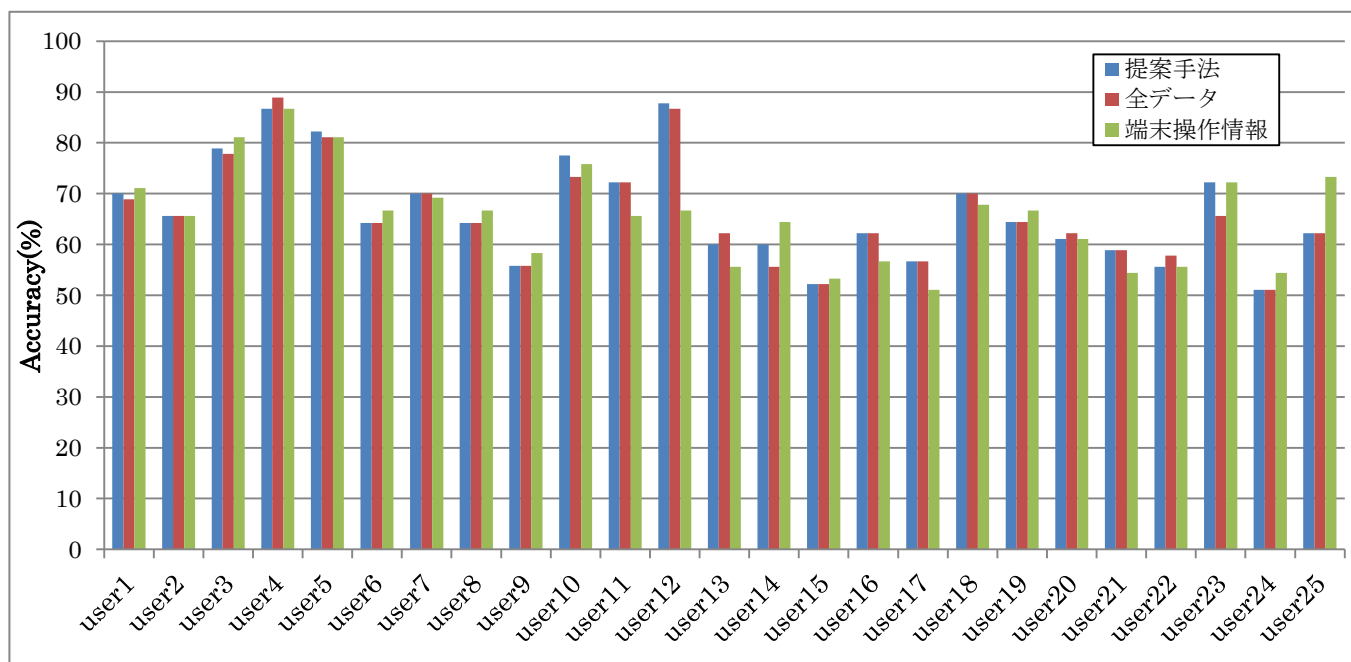


図 12 学習者毎の Accuracy (自信があるか否か)

## c) 個人差に関する実験の評価

個人差に関して考察すべきは、

- ・ 提案手法の精度に個人差があるか
- ・ 各手法を適用したときの結果の差分に個人差があるか

であるとする。以下、この 2 点に関し、「悩んでいるか否か」の推定、「自信があるか否か」の推定に分けて考察を行う。

### c-1) 悩んでいるか否かの推定における個人差の評価

#### c-1-1) 全体評価

図 11 において、「悩んでいるか否か」推定時の学習者ごとの Accuracy を見ると、提案手法の推定精度には個人差が見られる。user20 と user21 では、Accuracy の値に 30 ほど差がある。一方で、「悩んでいるか否か」推定時の各手法の推定精度の差分には、大きな個人差は見られない。図 9、図 10 で各手法と提案手法の差分を見ても、大きな値をとっていないことが見て取れる。すなわち、

- ・ 「悩んでいるか否か」の推定精度には個人差がある
- ・ 「悩んでいるか否か」の推定に各手法を用いた時の結果の差分には大きな個人差は無い

ということが分かる。

#### c-1-2) 各手法を適用したときの結果の差分に個人差が無いことに関する考察

図 8 において「悩んでいるか否か」推定時の各特微量の重み付けを見てみると、端末操作情報に対する重み付けが高く、表情・動作情報に対する重み付けは低い。ゆえに、必然的に、端末操作情報のみを用いた場合と、表情・動作情報を用いた場合の性能に大きな差が生まれなくなる。各手法を用いた時の結果の差分に個人差が見られない原因の一つは、この重み付け結果にあると考えられる。

また、推定がうまくいっていないユーザに関して、『「悩んでいるか否か」が表出する特微量が他のユーザと異なる』という仮説が立てられる。提案手法では学習者毎に特微量やその重みを変えるような処理をしていないため、もし表情・動作情報に「悩んでいるか否か」の状態が表出しやすい場合、提案手法では十分な精度が出せない。しかし、図 11 で推定が特に上手くいっていないユーザ（user7, user21）を見ると、ともに表情・動作情報を使用した時よりも、端末総情報のみを用いた時のほうが高い推定結果が出ている。つまり、今回の実験で「悩んでいるか否か」の推定がうまくいかなかったユーザは、提案手法で使用している

特徴量すべてに「悩んでいるか否か」の状態が表出しにくいのであると考えられる。

以上より、さらなる精度の向上を目指すには、より多くの特徴量を使用し、今回の特徴量に「悩んでいるか否か」の状態が表出しなかったユーザにも有効な特徴量を見出す必要がある。

## **c-2) 自信を持って答えたか否かの推定における個人差の評価**

### **c-2-1) 提案手法の精度に個人差があるか**

図 12 において、「自信があるか否か」推定時の学習者ごとの **Accuracy** を見ると、「悩んでいるか否か」同様、提案手法の推定精度には大きな個人差がある。原因としては、提案手法で使用している特徴量すべてに学習状態が表出しにくいユーザがいたこと、学習者毎に学習状態が表出する特徴量に違いがあったと考えられる。

### **c-2-2) 各手法を適用したときの結果の差分に個人差があるか**

図 12 の **user12** を見ると、端末操作情報のみを用いた場合には推定精度が低い、表情・動作情報を加える事で推定精度が大きく向上しているのが見て取れる。すなわち、**user12** は「自信があるか否か」の状態が、表情・動作情報に表出しやすいユーザであると言える。提案手法による推定も高い精度を示しており、提案手法が上手く機能したケースである。一方で、**user25** を見てみると、端末操作情報のみを用いた場合よりも、提案手法を用いた場合のほうが低い **Accuracy** をとってしまっている。これは、**user25** が、「自信があるか否か」の状態が、端末操作情報に表出しやすいユーザであったためと考えられる。以上から、「自信があるか否か」の状態が表出する特徴量には個人差があると考えられる。

### **c-2-3) 精度向上に関する考察**

学習状態が表出する特徴量に個人差があるということは、その学習者に特化した識別器、あるいは、傾向の似たユーザで学習された識別器を使用することで精度の向上を図れるということである。2010 年中村ら[3]の研究では、個人毎に識別器を用意することで高い品質の学習状態推定を行なっている。しかし、個人毎に識別器を用意するには、事前に各ユーザの正解セットが必要であり、状況としては一般的とはいえない。2009 年繁田ら[2]の研究のように、傾向に応じてユーザをグループ分けし、傾向の似たユーザで学習・識別することにより精度を向上する方法も考えられる。しかし、こちらの手法も、学習者がどのグループに属するのが事前に分かっている必要があり、一般的な状況とはいえない。これは、「環境依存度」と「精度・再現率」のトレードオフとも言える問題である。提案手法では「環境依存度の低いシステム」を前提としているため、表情・動作情報が有効なユーザの傾向を機械的に抽出するなどし、事前データ無しに個人間の差異を考慮した推定を行うようなシステムが求められる。

### 4.3.5 傾向の似たユーザで学習を行う手法の検討

#### a) 実験概要

4.3.4では個人差に関して実験・評価を行った。実験の結果、“自信があるか否か”の推定では、同じ被験者でも、

- ① 提案手法
- ② 全データを使用し、MKLを用いて分類する手法
- ③ 端末操作情報のみを利用し、MKLを用いて分類する手法

の推定結果に大きな差が見られるケースがあった。

この結果から、傾向の似た被験者で学習した結果を用いることで、精度の向上を図れるのではないかと考え、実験を行った。

実験では、被験者 20 名を、

- ・ 全てのデータを使用した場合のほうが Accuracy が高かったグループ
- ・ 端末操作情報のみを用いた場合のほうが Accuracy が高かったグループ

の 2 つのグループに分けた。全データを使用した場合の方が Accuracy が高いグループは、表情・動作に学習状態が表出しやすいグループであると言える。なお、全 25 名の被験者のうち、全てのデータを用いたときの結果と端末操作情報のみを用いたときの結果に大きな差が見られない被験者 5 名分のデータは除いた。グルーピングの結果を以下に示す。なお、ユーザ番号は図 11、図 12 に示したユーザ番号と対応している。

悩んでいるか否かの推定におけるグルーピング

全てのデータを使用したほうが Accuracy が高いグループ		端末操作情報のみのほうが Accuracy が高いグループ	
User6	User16	User1	User13
User11	User17	User4	User19
User12	User18	User7	User20
User14	User23	User8	User21
User15	User24	User9	User25

## 自信があるか否かの推定におけるグルーピング

全てのデータを使用したほうが Accuracy が高いグループ		端末操作情報のみのほうが Accuracy が高いグループ	
User4	User16	User1	User15
User7	User17	User6	User19
User11	User18	User8	User23
User12	User21	User9	User24
User13	User22	User14	User25

以上のようにグルーピングし、さらにグループ内で 5 名ずつの 2 グループに分ける。そして、片方のグループを学習セット、残りのグループをテストセットして推定を行う。その後、テストセットとして使用したグループで学習し、学習セットとして使用したグループをテストセットとして推定を行う。なお、提案手法では「全てのデータを使用した場合のほうが Accuracy が高かったグループ」のデータで実験を行った。こうして、各手法 10 名分のデータに対し学習状態の推定を実行した。

### b) 結果

“悩んでいるか否か”，“自信があるか否か” の推定を行い、各手法の Accuracy を算出した。結果を図 13 に示す。

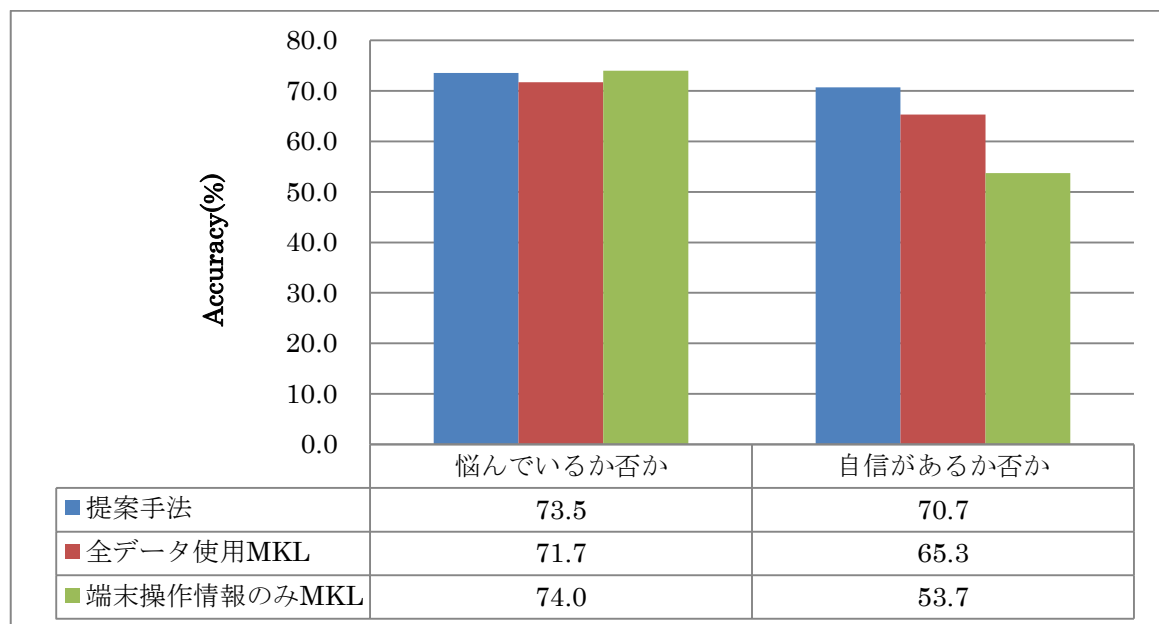


図 13 傾向の似たグループで学習を行った時の実験結果 (Accuracy)

さらに、提案手法と他の手法との間に統計的有意差があるのかを調べるため、マン・ホイットニーの U 検定を行った。結果を表 13 に示す。なお、網掛けは、その値が 0.05 以下であることを表す。

表 13 提案手法と各手法の Accuracy の差に関する検定結果 (P 値)

	悩んでいるか否か	自信があるか否か
全データ使用	0.28	0.23
端末操作情報使用	0.50	0.0086

## c) 評価

### c-1) 悩んでいるか否かの推定結果に関する評価

実験結果 図 13 を見ると、“悩んでいるか否か”の推定では、各手法の結果に大きな差は見られず、図 7 の実験結果と値も大きくは変わらない。検定結果 表 13 を見ても、各手法の間に統計的有意差は見られない。4.3.4 では、“悩んでいるか否か”推定時の表情・動作情報に対する重み付けが低いことについて言及した。今回の実験では、表情・動作情報に学習状態が表出しやすい被験者に対して性能が向上するような手法を検討しており、表情・動作情報が大きな意味を持たない“悩んでいるか否か”の推定で大きな結果の変化が生じなかったのは、結果として妥当であると言える。

### c-2) 自信があるか否かの推定結果に関する評価

実験結果 図 13 を見ると、“自信があるか否か”の推定では、各手法の結果に差が見られ、図 7 の実験結果と比べると、提案手法の結果は向上し、端末操作情報のみを用いた手法の結果は値が低くなっている。

4.3.4 では、“自信があるか否か”の推定において、表情・動作情報に学習状態が表出しやすいユーザと、表出しにくいユーザが存在することについて述べた。ゆえに、今回の実験における「傾向の似たユーザによる学習」が有効に働き、提案手法の Accuracy が向上したものと考えられる。一方で、端末操作情報のみを用いた手法の結果が低くなった理由としては、グループ分けをした結果学習データが減ってしまったことが考えられる。

検定結果としては、提案手法の「端末操作情報のみを利用し、MKL を用いて分類する手法」に対する優位性のみが、有意水準 5%において示された。「全データを使用し、MKL を用いて分類する手法」と提案手法の比較では、グルーピングを行わなかった場合 (図 7)、行

った場合（図 13），ともに提案手法が高い数値を記録しているが，統計的に有意な差は示すことができなかった．

#### 4. 3. 6 学習状態の表出が弱いデータを除去した実験

##### a) 実験概要

本研究では，正解セットとしてアンケートの結果を用いている．アンケートでは，各問題に対して以下の様な質問が与えられ，被験者は自分が問題に解答した時のことを思い出してこの質問に答える．

【問 XX】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 XX】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

しかし，アンケート結果は被験者の記憶や性格に左右されてしまい，必ずしも正確なデータが取れるとは限らない．特に，「少し悩んだ」，「あまり悩まなかった」，「あまり自信がなかった」，「少し自信があった」といった中間に位置する回答は，被験者の正確な感情では無い可能性や，被験者の感情が表情や端末操作情報に表出しにくい可能性が考えられる．

そこで本実験では，アンケートへの回答が「少し悩んだ」，「あまり悩まなかった」，「あまり自信がなかった」，「少し自信があった」のいずれかに当てはまるデータを除くことで，システムの精度が向上するか否かを調査する．

## b) 実験結果

アンケートへの回答が「少し悩んだ」、「あまり悩まなかった」、「あまり自信がなかった」、「少し自信があった」のいずれかに当てはまるデータを除いた結果、

悩んでいるか否か : 750 データ (25 人×30 問) → 351 データ

自信があるか否か : 750 データ (25 人×30 問) → 384 データ

となり、データ数は約半分となった。

上記の通りに絞り込みを行ったデータを使用し、学習・テストを行った。実験では、4.3.2 と同様に、全ユーザを 5 つのグループに分けて leave-one-out cross-validation を行った。Accuracy の算出結果を図 14 に示す。なお、「4.3.2 における提案手法」として示したデータは、学習データ、テストデータ、ともに全てのデータを使用したものである。

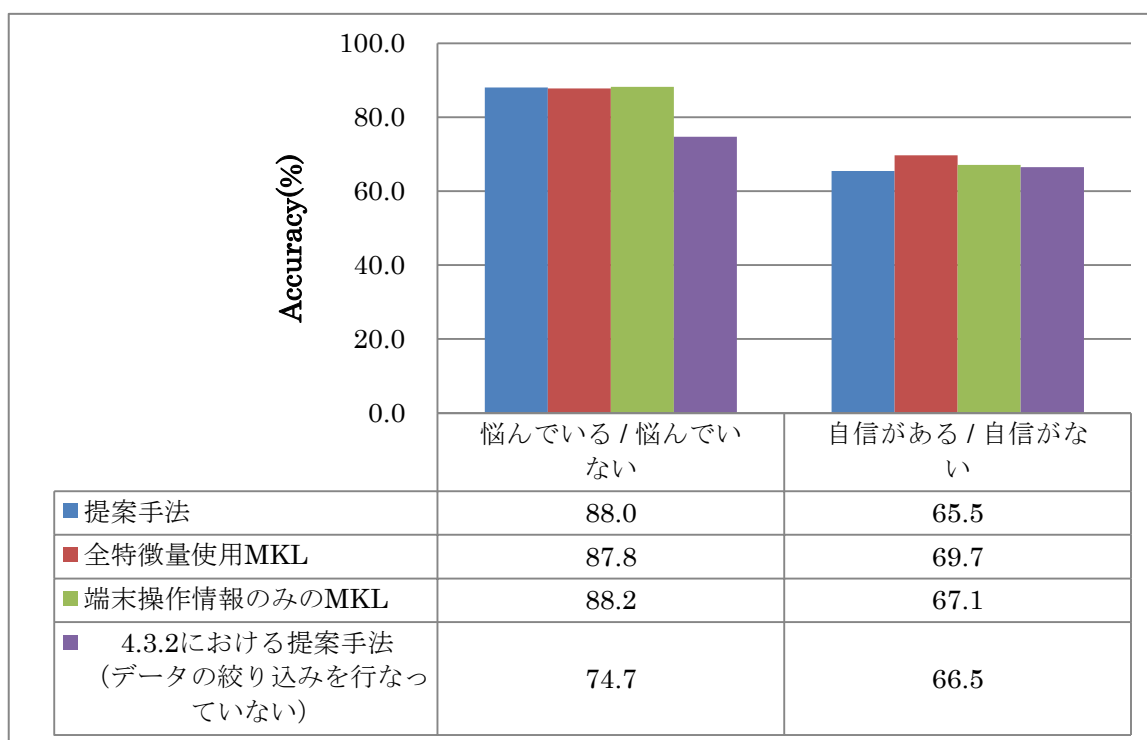


図 14 学習データ・テストデータを制限したときの実験結果 (Accuracy)

実験結果から、「悩んでいるか否か」推定時の Accuracy が、データの絞り込みを行っていない 4.3.2 の結果よりも高くなっているのが見て取れる。実験結果の差が有意なものかを確かめるため、マン・ホイットニーの U 検定を行った。結果を表 14 に示す。なお、網掛け



は 0.05 を下回る値を示す。検定の結果、データの絞り込みを行ったときと行わなかった時の提案手法の Accuracy には、統計的に有意な差があることが分かった。学習データ・テストデータを、学習状態の表出が強い物に絞り込むことで、高い精度の検出が行えることが分かった。

表 14 学習データ・テストデータを制限したときの実験に対する検定結果 (P 値)

	悩んでいるか否か	自信があるか否か
全データ使用 — 提案手法	4.69E-01	1.24E-01
端末操作情報使用 — 提案手法	4.76E-01	3.54E-01
4.3.2 における提案手法 — 提案手法	3.45E-07	4.44E-01

さらに、学習状態の表出が強いデータのみで学習を行った識別器で、アンケートへの回答が「少し悩んだ」、「あまり悩まなかった」、「あまり自信がなかった」、「少し自信があった」のいずれかに当てはまるデータを含む全てのデータを識別した。Accuracy の算出結果を図 15 に示す。

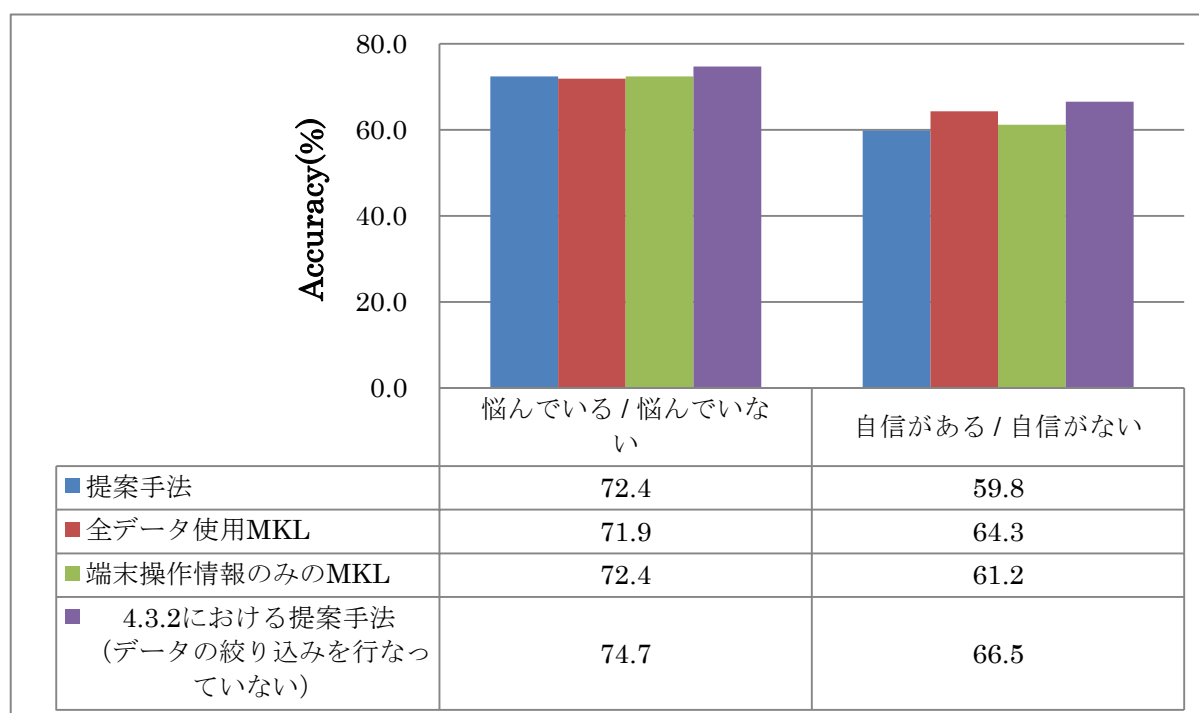


図 15 学習データのみを制限したときの実験結果 (Accuracy)

テストデータを全てのデータとした場合には、学習状態の表出が強い場合のデータのみを識別した時とは異なり、データを絞り込んで学習した識別器に優位性は見られなかった。

また、学習データを全てのデータとし、テストデータから、アンケートへの回答が「少し悩んだ」、「あまり悩まなかった」、「あまり自信がなかった」、「少し自信があった」のいずれかに当てはまるデータを取り除いた場合には、提案手法の **Accuracy** は 87.1%であった。

### c) 評価

学習データ、テストデータから学習状態の表出が弱いと考えられるものを除外することにより、「悩んでいるか否か」の推定精度が向上した。「悩んでいるか否か」の推定は、4.3.2 から 4.3.5 の実験でも高い精度を記録しており、本研究で使用している特徴量に強く学習状態が表出していると考えられる。テストセットを全データとすると 4.3.2 と同程度の **Accuracy** となることや、テストデータのみを絞り込んだ時にも高い **Accuracy** を記録したことから、「学習データを絞り込むことによって精度が上がる」というわけではなく、「学習状態の表出が強い時は高い精度で推定が行えている」のだと考えられる。

「自信があるか否か」に関しては精度の向上が見られなかった。「自信があるか否か」の推定では、“本研究で使用している特徴量に学習状態が表出しにくいユーザの存在”，“ユーザ毎に学習状態が表出する特徴量が異なる可能性”が考えられ、データを絞り込んでもこれらの問題を解決するには至らなかったものと考えられる。

## 4.4 総合評価

提案手法における新規点は、以下の4点である。

- ① タブレットデバイスを用いた学習状態の推定
- ② 特殊な装置を用いない
- ③ 学習教材に対する制限が弱い
- ④ 表情・動作が安定して取得できないことを考慮した学習状態推定

また、表情変化の個人差を考慮した既存手法が存在するのに対し、我々の手法では個人差を考慮せずに推定を行なっている。

したがって、今回の実験で評価すべきは、

- ・ タブレットデバイスのみを用いた、特殊な学習環境下でない学習状態の推定で、十分な精度・再現率を達成出来るか
- ・ 表情・動作が取得できないことを考慮した手法により、精度・再現率が向上するか
- ・ 推定結果に個人差が出ているか

の3点であると考える。以下で、この3つの観点から提案手法の評価を行う。

#### 4.4.1 タブレットデバイスのみを用いた学習状態推定に関する評価

##### a) 既存研究との比較

既存研究では、特殊な学習環境下において、高い精度で学習者の感情推定をおこなっている。例えば、表情・動作情報を用いた研究としては、2010 年中村ら[3]の研究で、ステレオカメラを使用し、個人に特化した識別器を生成した環境下で、学習者が「難しい」と感じている状態を 84.6%の精度で推定している。2002 年中村ら[5]の研究では、独自の学習ソフト上で、操作時間間隔のみから、約 8 割の再現率で行き詰まり状態を検知している。

図 4、図 5 を見ると、提案手法における精度・再現率は、既存手法と比べると低い値をとっている。しかしながら、「悩んでいない」状態に関しては、精度 77.4%，再現率 74.4%という値をとっており、平均値で見た時のシステムの品質としては低くないと考えられる。特に、「とても悩んだ」、「まったく悩まなかった」状態に関しては、Accuracy：87.1%という高い精度での推定に成功している。図 11 で「悩んでいるか否か」推定時の学習者ごとの Accuracy を見ても、個人差はあるものの、多くのユーザに対して 70 を超える高い値を記録しており、「悩んでいるか否か」の推定に関しては高い品質の推定を行えている。

##### b) “悩んでいるか否か”の推定で高い精度を出せた理由に関する考察

提案手法では、タッチデバイス特有の情報を使用すると同時に、MKL により各特徴量に最適な重みを学習している。環境を制限しないにも関わらず高い精度・再現率を実現できたのは、これらが効率的に機能したからであると考えられる。実際に、図 8 の「悩んでいるか否か」推定時の重み付け結果を見ると、タッチデバイス特有の特徴量である「解答を選んだ時のタッチ時間」や「タッチ時間の合計」が、「目の動き」、「眉毛の動き」など、既存手法で有用性が明らかになっている特徴量と同程度に重み付けされている。また、図 4、図 5、図 6 を見ると、大半の項目で、重み付けなしの推定結果を提案手法の推定結果が上回っている。この結果は MKL による特徴量の最適な重みの学習が機能していることを示している。

##### c) “自信があるか否か”の推定で精度が低かった理由に関する考察

「自信があるか否か」の推定では、平均 Accuracy：66.5%（図 7）と、あまり高い数値は出していない。図 12 を見ると、学習者ごとの推定精度にも大きな差があり、安定した推定が行えているとは言いがたい。「悩んでいるか否か」の推定と比べて「自信があるか否か」の推定性能が劣っている原因としては、今回使用した特徴量が「自信があるか否か」の推定に適切でなかった可能性が挙げられる。推定精度を上げるには、多様な特徴量を使用し、その中から「自信があるか否か」の推定に有用な特徴量を探し出す必要がある。

#### d) タブレットデバイスのみを用いた学習状態推定に関する総評

以上の結果から、環境を限定している既存手法ほどの精度は出ないものの、「悩んでいるか否か」の推定に関しては、タッチデバイスから得られる情報のみで高い品質の学習状態推定を行えることが分かった。また、「自信があるか否か」の推定に関しては課題を残す結果となっており、より多様な特徴量の使用で精度を向上できる可能性がある。

### 4. 4. 2 表情・動作が安定して取得できないことを考慮した手法に関する評価

図 6 を見ると、「悩んでいる」、「自信がある」、「自信がない」状態の推定において、すべてのデータに対して単純に MKL を適用した場合よりも、提案手法のほうが高い F 値をとっていることが分かる。図 7 を見ると、Accuracy においても提案手法が他の手法を上回っている。しかし、統計的に提案手法の優位性が示されたのは、重み付けをせずに全データを使用した場合に対する Precision, F 値のみであった。

“悩んだか否か”の推定に関しては、表情・動作情報が推定において大きな意味を持たず、提案手法における「表情・動作情報の取得状態による使用特徴量の変更」が与える影響も無いに等しかった。一方で、“自信があるか否か”の推定では、表情・動作情報が推定に有用な情報であるにもかかわらず、提案手法と他の手法に統計的に有意な差が示されなかった。これは、表情・動作情報に学習状態が表出しやすい被験者とそうでない被験者がいるにもかかわらず、特徴量の重み付け全被験者に対して同様に行ったことが原因であると考えられる。実際に、4. 3. 5 では、傾向の似たユーザで学習したデータを用いることで、提案手法が、端末操作情報のみを使用する手法に対して統計的に有意な差を示した。また、統計的に有意ではなかったものの、「全データを使用し、MKL を用いて分類する手法」に対しても約 5% の差をつけた。表情・動作情報が有効に働くという結果が出ている以上、提案手法により正確でないデータを除外することは意味があると考えられる。より多くのデータによる実験、より細かなグルーピング、特徴量の追加などにより、統計的に有意な差を示せるだけの手法を提案することが、今後の課題である。

### 4. 4. 3 個人差に関する評価

4. 3. 4 では、個人差に関する実験・評価を行った。“悩んでいるか否か”の推定、“自信があるか否か”の推定、ともに提案手法による推定精度には個人差が見られた。これには、

- ・ 今回使用した特徴量の中に学習状態が表出する特徴量が存在しないユーザがいる
- ・ 学習状態が表出する特徴量が低く重み付けされているユーザがいる

という原因が考えられる。実際に、4. 3. 5 における「傾向の似たユーザで学習を行う手法」では、提案手法の精度の向上が見られた。これは、学習状態が表出する特徴量が低く重み付

けされていたユーザがいることを示す.

傾向の似たユーザで学習を行うことにより精度向上に成功したが,

- ・ より細かな傾向の調査・グルーピング
- ・ 新たな特徴量の追加

などにより, 更に推定精度を向上できる可能性がある.

しかしながら, 事前にユーザの学習状態表出傾向が分かっているような状態は一般的ではない. より一般的な環境下で高い推定精度を目指すには, 表情の変化の大小などから機械的にユーザの学習状態表出傾向を識別する技術が必要となる.

## 4.5 実験・評価のまとめ

最後に, 本研究で行った実験・評価, 及び得られた知見について, 表 15 にまとめる

表 15 実験・評価・得られた知見

評価項目	実験結果・評価・得られた知見
提案手法の性能 (Accuracy)	<u>悩んでいるか否か : 74.7%</u> <u>自信があるか否か : 66.5%</u>
他の手法との比較	▼実験結果 「表情・動作の取得状態を考慮せず, 全特徴量を用いて MKL を使用する手法」, 「端末操作情報のみを用いて MKL を使用する手法」と比較した結果, <u>統計的に有意な差は確認されなかった.</u>  ▼考察 端末操作情報のみで高い精度の推定が行える. 表情・動作情報は端末操作情報と比べると重み付けも低く, 提案手法の効果も見受けられなかった.
効果的であった特徴量	○悩んでいるか否か 解答時間, 解答を選んでから答えるまでの時間, 特徴点検出成功フレームの割合

	<p>○自信をもって答えたか否か</p> <p>解答時間，目の動き，解答を選んだ時のタッチ時間</p>
個人間の差異	<p>▼実験結果</p> <p>「悩んでいるか否か」の推定，「自信があるか否か」の推定，ともに<u>提案手法の推定精度には個人差が見られた</u>。</p> <p>「自信があるか否か」の推定では，表情・動作情報を使用した方が大きく精度が向上する被験者がいた。逆に大きく精度が低下する被験者もいた。</p> <p>▼考察</p> <p><u>今回使用した特徴量に学習状態が表出する被験者と表出しない被験者がいたのだと考えられる。また，個人毎に「自信があるか否か」の学習状態が表出する特徴量には違いがあると考えられる。</u></p>
傾向の似たユーザをグルーピングして学習・テストした場合の性能	<p>▼実験結果</p> <ul style="list-style-type: none"> <li>・悩んでいるか否か</li> </ul> <p>提案手法とその他の手法に差は見られなかった</p> <ul style="list-style-type: none"> <li>・自信を持って答えたか否か</li> </ul> <p>端末操作情報のみを使用した場合と比べて，提案手法が統計的に有意な差で優れていた</p> <p>▼考察</p> <p><u>「自信を持って答えたか否か」の推定では，表情・動作に学習状態が表出しやすいユーザと表出しにくいユーザが存在しており，グルーピングが有効に働いた。</u></p>
学習状態の表出が大きいデータに絞った場合の推定性能	<p>▼実験結果</p> <ul style="list-style-type: none"> <li>・悩んでいるか否か</li> </ul> <p>Accuracy88.0%と，大きく性能が向上した。</p> <p>学習を全てのデータで行い，テストデータを学習状態の表出が大きいものに絞った場合も，<u>Accuracy87.1%を記録した</u>。</p>

	<ul style="list-style-type: none"> <li>・自信を持って答えたか否か</li> </ul> <p>Accuracy に大きな変化は見られなかった.</p> <p>▼考察</p> <p>アンケートで中間的な回答（「あまり〇〇，少し〇〇等」）を得たデータは，必ずしも正確に学習状態を表しているとは言えない.</p> <p><u>「とても悩んだ／」まったく悩まなかった」という状態の推定に絞って考えれば，非常に高い精度で推定を行えていることが分かった.</u></p>
--	---

## 第5章 おわりに

### 5.1 本研究のまとめ

本稿では、一般的な学習環境下で生徒の学習状態を機械的に理解することを目的として、タブレットデバイスから得られる表情・動作情報と端末操作情報を統合し、学習状態の推定を実現する手法を提案した。実験の結果、“解答時に悩んだか否か”、“自信があったか否か”の推定について、それぞれ平均精度 75.3%, 66.7%, 平均再現率 74.8%, 60.1%を達成し、タブレット端末特有の情報の利用、MKLの使用、表情・動作取得状態に応じた特徴量選択により、一般的な環境下でも高い精度で推定を行えることを示した。特に、「とても悩んだ」、「まったく悩まなかった」状態に関しては、Accuracy : 87.1%という高い精度での推定に成功している。さらに、傾向の似たユーザで学習したデータを使用することで、精度の向上が見込めることを、実験により示した。

### 5.2 今後の課題

今後の課題としては、より環境依存度の低い手法の提案が挙げられる。選択問題・英語問題以外の学習環境や、閲覧主体型の学習環境に対しても提案手法を適用し、改善を加え、より学習環境に対し柔軟に対応できるシステムを目指したい。

また、第二の課題として、使用特徴量に関する検討が挙げられる。今回取得した特徴量は、既存手法を基に自分で定義したものだが、タブレットデバイスを用いて取得できるデータはまだ残されている。加速度センサ、圧力センサなど、タブレットデバイスに搭載されたセンサを活用することで、より精度の高いシステムを構築できる可能性がある。

さらに第三の課題として、個人差を考慮したシステムの構築が挙げられる。本研究では、「自信があるか否か」の推定における表情・動作情報の使用に関し、個人間で特徴量の重要度に差異があることが示された。この差異を事前データ無しに検出し、特徴量の重み付けに適用することができれば、より質の高いシステムを構築することが可能であると考えられる。



## 謝 辞

本研究を行うにあたり，数々のご指導を頂いた山名早人教授に厚くお礼申し上げます．

また，日々の研究活動の中で様々なアドバイスをくださった研究室の皆様に深く感謝致します．

## 研究業績

須藤 優介，山名 早人：“表情・動作情報とタッチパネル操作情報の統合による学習状態の推定”，DEIM2013.（投稿中）

## 参考文献

- [1] 黒川 智哉, 野須 潔, 山崎 清之, “顔画像および生体信号のテンプレートを用いた e-Learning 学習者の心理状態時系列変化推定の検討”, 映像情報メディア学会誌, Vol.61, No.12, pp.1779-1784, 2007.
- [2] 繁田 亜友子, 小池 武士, 黒川 智哉, 野須 潔, “e-Learning 学習者の顔画像解析による感情推定アルゴリズムの研究”, 電気学会論文誌 C, Vol.129, No.10, pp.1881-1888, 2009.
- [3] 中村和晃, 角所考, 村上正行, 美濃導彦, “e-learning における学習者の顔動作観測に基づく主観的難易度の推定”, 電子情報通信学会論文誌 D, Vol.J93-D, No.5, 2010.
- [4] D. Szafir and B Mutlu , “Pay attention! : designing adaptive agents that monitor and improve user engagement”, CHI 2012, pp.11-20, 2012.
- [5] 中村喜宏, 赤松則夫, 桑原恒夫, 玉城幹介, “操作時間間隔の変動に着目した CAI 学習の行き詰まり検知方法”, 電子情報通信学会論文誌, Vol.J85-D-I, No.1, pp.79-90, 2002.
- [6] A. Kapoor and R. W. Picard , “Multimodal affect recognition in learning environments”, MULTIMEDIA 2005, pp.677-682, 2005.
- [7] 中村 和晃, 角所 考, 美濃 導彦, “e-learning 中の学習者の端末操作と顔画像から推定される主観的難易度を用いた教材評価の試み”, 情報処理学会研究報告 第 5 回 CMS 研究会, pp.41-48, 2007.
- [8] 総務省, “フューチャースクール推進事業”,  
[http://www.soumu.go.jp/main\\_sosiki/joho\\_tsusin/kyouiku\\_joho-ka/future\\_school.html](http://www.soumu.go.jp/main_sosiki/joho_tsusin/kyouiku_joho-ka/future_school.html)  
(2013.1.5 アクセス)
- [9] T.F. Cootes, C.J. Taylor, D. Cooper, and J. Graham , "Active Shape Models -Their Training and Application", Computer Vision and Image Understanding, Vol. 61, No. 1, pp. 38-59, 1995.
- [10] D. McDuff, A. Karlson, A. Kapoor, A. Roseway and M. Czerwinski , “AffectAura: an intelligent system for emotional memory”, CHI 2012, pp.849-858, 2012.
- [11] “ASMLibrary”,  
<http://code.google.com/p/asmlibrary/>  
(2013.1.11 アクセス)

- [12] F.R. Bach, G.R.G. Lanckriet and M.I. Jordan, “Multiple kernel learning, conic duality, and the SMO algorithm”, International Conference on Machine Learning, p.6, 2004.

## 付録 A 出題した問題の例

### 【問 1】（センター試験レベル）

Some companies have (     ) a new policy of using English as the official in-house language.

1. absorbed
2. accompanied
3. adopted
4. appointed

### 【問 10】（センター試験レベル）

After he joined the travel agency, he worked hard to improve his English in order to carry (     ) his duties more effectively.

1. away
2. back
3. off
4. out

### 【問22】（中学レベル）

I'm looking for a job now.

訳：私は今新しい仕事を（ ？ ）.

1. しっかり見ている
2. あきあきしている
3. 探(さが)している
4. がんばっている

## 付録 B 問題解答後のアンケート

【問 1】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 1】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 2】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 2】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 3】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 3】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 4】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 4】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 5】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 5】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 6】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 6】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 7】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 7】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 8】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 8】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 9】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 9】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 10】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 10】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 11】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 11】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 12】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 12】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 13】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 13】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 14】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 14】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 15】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 15】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 16】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 16】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 17】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 17】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 18】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 18】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった



【問 19】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 19】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 20】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 20】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 21】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 21】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 22】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 22】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 23】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 23】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 24】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 24】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 25】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 25】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 26】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 26】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 27】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 27】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 28】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 28】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 29】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 29】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

【問 30】で悩んだか否か

- ☐ とても悩んだ
- ☐ 少し悩んだ
- ☐ あまり悩まなかった
- ☐ まったく悩まなかった

【問 30】で自信を持って答えたか否か

- ☐ まったく自信がなかった
- ☐ あまり自信がなかった
- ☐ 少し自信があった
- ☐ かなり自信があった

## 付録 C 個人毎・問題毎のアンケート集計結果, 正答率

Table 1, Figure 1 に, 個人毎のアンケート集計結果・正答率, Table 2, Figure 2 に, 問題毎のアンケート集計結果・正答率を示す.

Table 1 個人毎のアンケート集計結果・正答率

	悩まなかった問題の割合(%)	自信があった問題の割合(%)	正答率(%)
user1	63.3	60.0	73.3
user2	46.7	66.7	50.0
user3	56.7	63.3	46.7
user4	66.7	90.0	96.7
user5	43.3	53.3	60.0
user6	60.0	40.0	56.7
user7	53.3	56.7	73.3
user8	53.3	70.0	83.3
user9	36.7	70.0	60.0
user10	26.7	23.3	50.0
user11	56.7	76.7	76.7
user12	73.3	93.3	96.7
user13	36.7	66.7	83.3
user14	33.3	50.0	70.0
user15	50.0	63.3	70.0
user16	70.0	70.0	63.3
user17	26.7	50.0	83.3
user18	66.7	73.3	76.7
user19	43.3	46.7	56.7
user20	83.3	80.0	83.3
user21	63.3	66.7	66.7
user22	73.3	70.0	56.7
user23	60.0	60.0	56.7
user24	53.3	73.3	86.7
user25	30.0	33.3	53.3

Table 2 問題毎のアンケート集計結果・正答率

	悩んだユーザの割合 (%)	自信がなかったユーザの割合 (%)	正答率 (%)
問 1	76.0	56.0	76.0
問 2	68.0	64.0	36.0
問 3	60.0	52.0	88.0
問 4	88.0	84.0	32.0
問 5	92.0	80.0	48.0
問 6	84.0	76.0	48.0
問 7	36.0	32.0	52.0
問 8	56.0	52.0	56.0
問 9	36.0	28.0	52.0
問 10	56.0	44.0	96.0
問 11	60.0	32.0	60.0
問 12	44.0	24.0	60.0
問 13	52.0	36.0	72.0
問 14	48.0	40.0	84.0
問 15	68.0	64.0	48.0
問 16	52.0	48.0	84.0
問 17	36.0	16.0	88.0
問 18	68.0	60.0	64.0
問 19	48.0	56.0	56.0
問 20	68.0	72.0	36.0
問 21	0.00	0.00	100
問 22	0.00	0.00	100
問 23	32.0	8.00	92.0
問 24	52.0	28.0	76.0
問 25	48.0	16.0	88.0
問 26	12.0	0.00	96.0
問 27	28.0	28.0	84.0
問 28	8.00	4.00	8.00
問 29	32.0	20.0	100
問 30	0.00	0.00	96.0

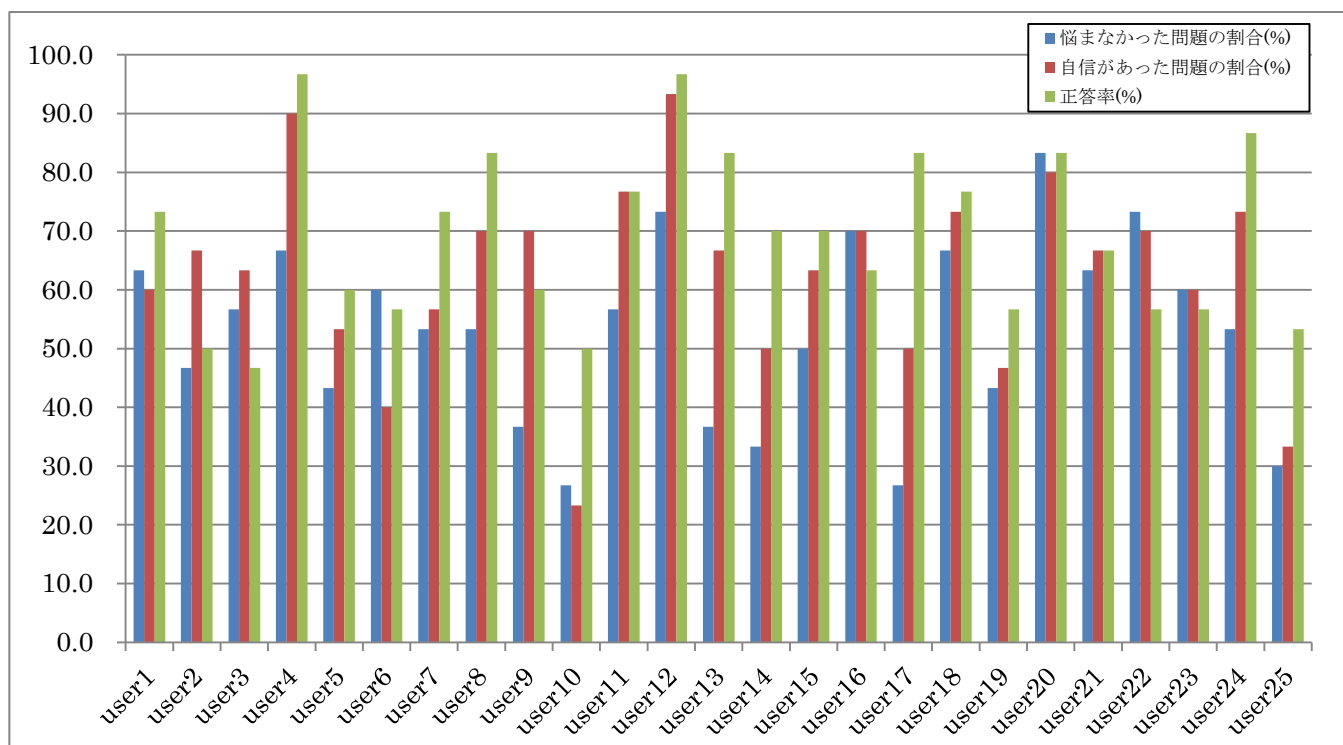


Figure 1 個人毎のアンケート集計結果・正答率

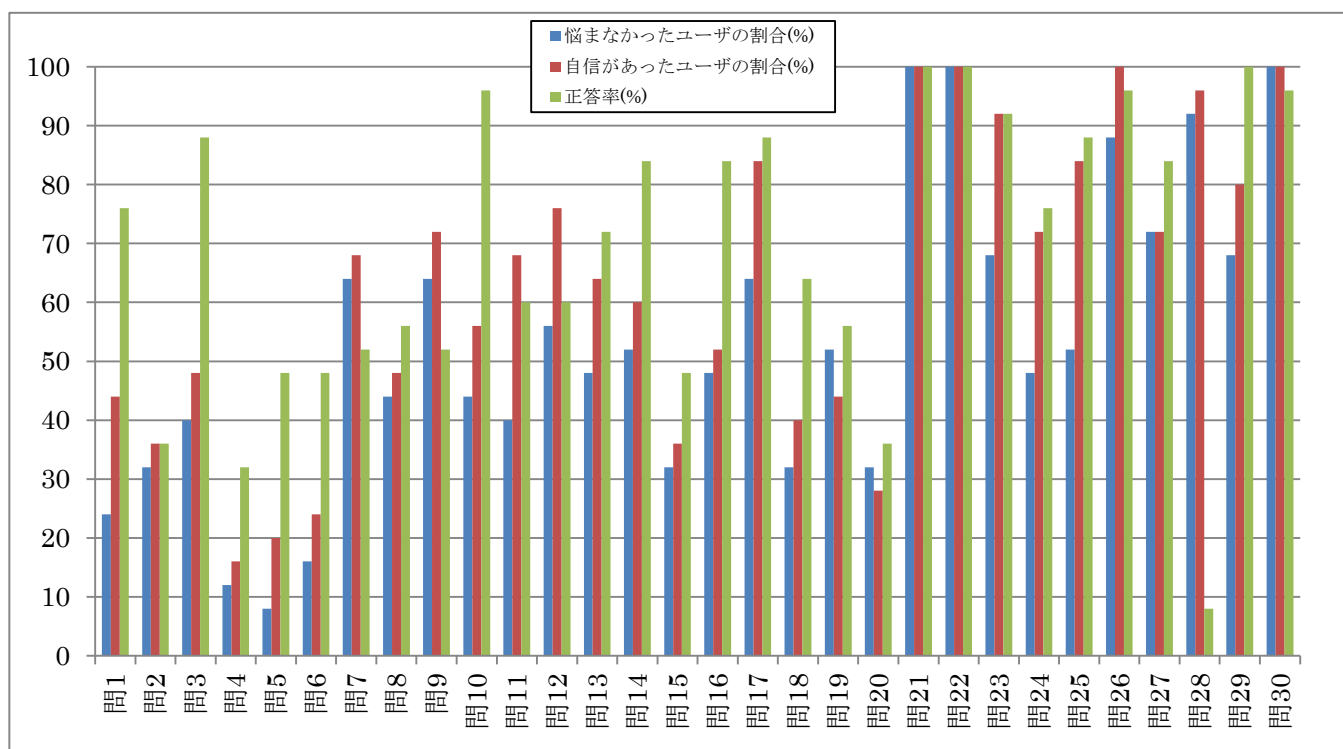


Figure 2 問題毎のアンケート集計結果・正答率